Cuu-Ho1525-19- Po264/

THE ACCOUNTING REVIEW

VOLUME 68, No. 1 JANUARY 1993

[illegible]

657.05
Ac 27

19

H. FRED MITTELSTAEDT and PHILIP R. REGIER
The Market Response to Pension Plan Terminations

RAJIV D. BANKER, SOMNATH DAS, and SRIKANT M. DATAR
Complementarity of Prior Accounting Information: The Case of Stock Dividend Announcements

JEFFREY L. CALLEN and HAIM FALK
Agency and Efficiency in Nonprofit Organizations: The Case of "Specific Health Focus" Charities

YAW M. MENSAH and SHU-HSING LI
Measuring Production Efficiency in a Not-for-Profit Setting: An Extension

ROBERT D. HAYES and JAMES A. MILLAR
A Rejoinder to "Measuring Production Efficiency in a Not-for-Profit Setting: An Extension"

MICHAEL ALLES and RUSSELL LUNDHOLM
On the Optimality of Public Signals in the Presence of Private Information

MOHINDER PARKASH and CAROL F. VENABLE
Auditee Incentives for Auditor Independence: The Case of Nonaudit Services

SHORTER ARTICLES

LARRY R. DAVIS, DAVID N. RICCHIUTE, and GREG TROMPETER
Audit Effort, Audit Fees, and the Provision of Nonaudit Services to Audit Clients

KENNETH S. LOREK, THOMAS F. SCHAEFER, and G. LEE WILLINGER
Time-Series Properties and Predictive Ability of Funds Flow Variables

FRANK A. BUCKLESS and ROBERT L. PEACE
The Influence of the Source of Professional Standards on Juror Decision Making

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education and research. Full membership entitles one to receive *The Accounting Review, Accounting Education News, Accounting Horizons*, and *Issues in Accounting Education*, and to purchase Association publications at a discount.

The members of the 1992–93 Executive Committee of the Association are as follows:

President	Gary L. Sundem, University of Washington
President-Elect	Andrew D. Bailey, Jr., University of Arizona
Vice President	Andre Zund, University of St. Gallen
Vice President	Gary K. Meek, Oklahoma State University
Vice President	Bernard J. Milano, KPMG Peat Marwick
Secretary-Treasurer	Jane O. Burns, Texas Tech University
Director of Education	Robert Ingram, University of Alabama
Director of Research	Katherine Schipper, University of Chicago
Director of Publications	Kenneth J. Euske, Naval Postgraduate School
Past President	Arthur R. Wyatt, University of Illinois at Urbana-Champaign

The Executive Director of the Association is Paul L. Gerhardt.

Membership in the Association is available at the following annual rates:

Full membership, US and Canada:	US $65.00
Full membership, other countries:	US $65.00
Associate membership:	US $12.00 for students residing in the United States and Canada

In addition, full members may belong to one or more Sections: Auditing; Information Systems/Management Advisory Services; Two-Year Colleges; International Accounting; Public Interest; Government and Nonprofit; Taxation; Administrators of Accounting Programs; Accounting, Behavior and Organizations; and Management Accounting. Each Section assesses its own annual dues which are collected by the Association.

The Accounting Review is published quarterly by the American Accounting Association—January, April, July, and October: printed by George Banta Company, Inc., Menasha, Wisconsin. Subscription price is $90.00 per year.

Second class postage paid at Sarasota, Florida and at additional mailing offices.

Copyright © American Accounting Association. All rights reserved. Printed in the United States of America. ISSN 0001-4826.

Editorial correspondence and all manuscripts should be sent to Professor A. Rashad Abdel-khalik, Editor, *The Accounting Review*, Fisher School of Accounting, College of Business Administration, University of Florida, Gainesville, FL 32611.

Effective June 1993, the editoral responsibility will be assumed by Professor Robert P. Magee of Northwestern Universtity.

The submission fee of $50.00 for members of the AAA or $100.00 for nonmembers must accompany manuscripts when submitted. Books for review should be sent to Professor Shane R. Moriarity, School of Accounting, University of Oklahoma, Norman, OK 73019. Request for reprint permission, subscriptions, advertisements, and dues should be sent to Mr. Paul Gerhardt, American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review*.

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE

AMERICAN ACCOUNTING ASSOCIATION

Editor
A. Rashad Abdel-khalik

Associate Editors

Robert H. Ashton	Robert P. Magee (*Editor Elect*)
Victor L. Bernard	James C. McKeown
J. Harry Evans, III	Ira Solomon
John S. Hughes	Shyam Sunder

VOL. 68 JANUARY 1993 No. 1

H. FRED MITTELSTAEDT and PHILIP R. REGIER
The Market Response to Pension Plan Terminations 1

RAJIV D. BANKER, SOMNATH DAS, and SRIKANT M. DATAR
Complementarity of Prior Accounting Information: The Case of Stock Dividend Announcements 28

JEFFREY L. CALLEN and HAIM FALK
Agency and Efficiency in Nonprofit Organizations: The Case of "Specific Health Focus" Charities 48

YAW M. MENSAH and SHU-HSING LI
Measuring Production Efficiency in a Not-for-Profit Setting: An Extension 66

ROBERT D. HAYES and JAMES A. MILLAR
A Rejoinder to "Measuring Production Efficiency in a Not-for-Profit Setting: An Extension" 89

MICHAEL ALLES and RUSSELL LUNDHOLM
On the Optimality of Public Signals in the Presence of Private Information 93

MOHINDER PARKASH and CAROL F. VENABLE
Auditee Incentives for Auditor Independence: The Case of Nonaudit Services 113

SHORTER ARTICLES

LARRY R. DAVIS, DAVID N. RICCHIUTE, and GREG TROMPETER
Audit Effort, Audit Fees, and the Provision of Nonaudit Services to Audit Clients 135

KENNETH S. LOREK, THOMAS F. SCHAEFER, and G. LEE WILLINGER
Time-Series Properties and Predictive Ability of Funds Flow Variables 151

FRANK A. BUCKLESS and ROBERT L. PEACE
The Influence of the Source of Professional Standards on Juror Decision Making 164

Contents continued on Overleaf

COMMENTS

JANE F. MUTCHLER and KURT PANY
A Comment on "The Influence of the Source of Professional Standards on Juror Decision Making" **176**

L. PETER JENNERGREN and BERTIL NASLUND
A Comment on "Valuation of Executive Stock Options and the FASB Proposal" **179**

TAYLOR W. FOSTER III, PAUL R. KOOGLER, and DON VICKREY
Valuation of Executive Stock Options and the FASB Proposal: An Extension **184**

AMERICAN ACCOUNTING ASSOCIATION

Committees 1992–1993 .... **190**
Occasional Reviewers **201**

BOOK REVIEWS, Shane Moriarity, Editor

Chambers, ***Foundations of Accounting*** ....DALE A. BUCKMASTER **205**
Davia, Coggins, Wideman, and Kastantin, ***Management Accountant's Guide to Fraud Discovery and Control*** ....JAMES W. BRACKNER **206**
Donaldson, ***Business Ethics: A European Casebook*** ....MARY GREENAWALT **207**
Fitzgerald, Johnston, Brignall, Silvestro, and Voss, ***Performance Measurement in Service Businesses*** ....RONALD J. LEWIS **207**
Lauer, Peacock, and Graesser, Editors, ***Questions and Information Systems*** ....MARTHA M. EINING **208**
Schweitzer, Trossman, and Lawson, ***Break-Even Analysis: Basic Model, Variants, Extensions*** ....MAURICE L. HIRSCH, JR. **209**
Steele, ***Audit Risk and Audit Evidence: The Bayesian Approach to Statistical Auditing*** ....RICHARD A. GRIMLUND **210**
List of Books Received **211**
Editorial Policy and Style Information **213**

P6241

PAST EDITORS

1926–1929	William A. Paton	1968–1970	Charles H. Griffin
1929–1943	Eric L. Kohler	1971–1972	Eldon S. Hendriksen
1944–1947	A. C. Littleton	1973–1975	Thomas F. Keller
1948–1949	Robert L. Dixon	1976–1978	Don T. DeCoster
1950–1959	Frank P. Smith	1978–1983	Stephen A. Zeff
1960–1962	Robert K. Mautz	1983–1987	Gary L. Sundem
1963–1964	Lawrence L. Vance	1987–1989	William R. Kinney, Jr.
1965–1967	Wendell Trumbull		

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Editor
A. Rashad Abdel-khalik

Associate Editors
Robert H. Ashton
Victor L. Bernard
J. Harry Evans, III
John S. Hughes
Robert P. Magee (*Editor Elect*)
James C. McKeown
Ira Solomon

19

VOL. 68 APRIL 1993 No. 2

A FORUM ON THE EFFECTS OF VIOLATING DEBT COVENANTS

KEVIN C. W. CHEN and K. C. JOHN WEI
Creditors' Decisions to Waive Violations of Accounting-Based Debt Covenants 218
MESSOD D. BENEISH and ERIC PRESS
Costs of Technical Violation of Accounting-Based Debt Covenants 233
SAMIR M. EL-GAZZAR
Stock Market Effects of the Closeness to Debt Covenant Restrictions Resulting from Capitalization of Leases 258
MARY BETH MOHRMAN
Debt Contracts and FAS No. 19: A Test of the Debt Covenant Hypothesis 273
CLIFFORD W. SMITH, JR.
A Perspective on Accounting-Based Debt Covenant Violations 289

MARY YORK CHRIST
Evidence on the Nature of Audit Planning Problem Representations: An Examination of Auditor Free Recalls 304
STEPHEN C. HANSEN
Strategic Sampling, Physical Units Sampling, and Dollar Units Sampling 323
SIEW HONG TEOH and T. J. WONG
Perceived Auditor Quality and the Earnings Response Coefficient 346

SHORTER ARTICLES

LI-CHIN JENNIFER HO
Option Trading and the Relation Between Price and Earnings: A Cross-Sectional Analysis 368
NORMAN STRONG and MARTIN WALKER
The Explanatory Power of Earnings for Stock Returns 385
ALAN S. DUNK
The Effect of Budget Emphasis and Information Asymmetry on the Relation Between Budgetary Participation and Slack 400

Contents continued on Overleaf

COMMENTS

SCOTT K. JONES and LAWRENCE A. PONEMON
A Comment on "A Multidimensional Analysis of Selected Ethical Issues in Accounting" 411

STEVEN M. FLORY, THOMAS J. PHILLIPS, JR., R. ERIC REIDENBACH, and DONALD P. ROBIN
A Reply to "A Comment on 'A Multidimensional Analysis of Selected Ethical Issues in Accounting' " 417

BOOK REVIEWS, Shane Moriarity, Editor

Barnes, ***Value-for-Money Audit Evidence*** .... PAUL A. JANELL 422
Bavishi, ***Global Company Handbook: An Analysis of the Financial Performance of the World's Leading 7,500 Companies, 1992 Edition*** .. MICHAEL R. KINNEY 423
Boys and Freear, Editors, ***Accounting History 1976–1986: An Anthology*** .... HANS V. JOHNSON 424
Bruns, Editor, ***Performance Measurement, Evaluation, and Incentives*** .... MARVIN L. BOUILLON 425
Cooke and Kikuya, ***Financial Reporting in Japan: Regulation, Practice and Environment*** .... HELEN GERNON 425
Dean, Joye, and Blayney, ***Strategic Management Accounting Survey: Overhead Cost Allocation & Performance Evaluation Practices of Australian Manufacturers*** .... SHU S. LIAO 426
Enthoven, Sokolov, and Petrachkov, ***Doing Business in Russia and the Other Former Soviet Republics: Accounting and Joint Venture Issues*** .... TIMOTHY S. DOUPNIK 427
Hoogendoorn and Bijl, editors, ***Financial Reporting in the Nineties: Regulation and Innovation at a Crossroads*** .... ADOLF J. H. ENTHOVEN 428
Lorenson, ***Accounting for Liabilities*** .... EILEEN PEACOCK 430
McKinnon, and Bruns, ***The Information MOSAIC*** .... FRANK R. PROBST 431
Monden, ***Cost Management in the New Manufacturing Age: Innovation in the Japanese Automotive Industry*** .... HELEN M. TRAUGH 432
Rahman, ***The Australian Accounting Standards Review Board: The Establishment of its Participative Review Process*** .... W. T. WREGE 433
Rosenthal, ***Meta-Analytic Procedures for Social Research, Revised Edition*** .... SOON-YONG KWON 434
Spencer, ***The How & Why of a Tax on Wealth*** .... DALE BANDY 434
Wei, Chow, and Cooper, ***Accounting and Finance in China: A Review of Current Practice*** .... ROBERT A. LYON 436
List of Books Received 437
Editorial Policy and Style Information 439

PAST EDITORS

1926–1929	William A. Paton	1968–1970	Charles H. Griffin
1929–1943	Eric L. Kohler	1971–1972	Eldon S. Hendriksen
1944–1947	A. C. Littleton	1973–1975	Thomas F. Keller
1948–1949	Robert L. Dixon	1976–1978	Don T. DeCoster
1950–1959	Frank P. Smith	1978–1983	Stephen A. Zeff
1960–1962	Robert K. Mautz	1983–1987	Gary L. Sundem
1963–1964	Lawrence L. Vance	1987–1989	William R. Kinney, Jr.
1965–1967	Wendell Trumbull		

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Editor
A. Rashad Abdel-khalik

Associate Editors
Robert H. Ashton
Victor L. Bernard
J. Harry Evans, III
John S. Hughes
Robert P. Magee (*Editor Elect*)
James C. McKeown
Ira Solomon

VOL. 68 JULY 1993 No. 3

JEFFREY J. McMILLAN and RICHARD A. WHITE
Auditors' Belief Revisions and Evidence Search: The Effect of Hypothesis Frame, Confirmation Bias, and Professional Skepticism **443**

S. MARK YOUNG, JOSEPH FISHER, and TIM M. LINDQUIST
The Effects of Intergroup Competition and Intragroup Cooperation on Slack and Ouput in a Manufacturing Setting **466**

JON S. DAVIS and CHARLES W. SWENSON
Experimental Evidence on Tax Incentives and the Demand for Capital Investments **482**

RUSSELL LANGER and BARUCH LEV
The FASB's Policy of Extended Adoption for New Standards: An Examination of FAS No. 87 .... **515**

MASAKO N. DARROUGH
Disclosure Policy and Competition: Cournot vs. Bertrand **534**

A FORUM ON COST DRIVERS

YAIR M. BABAD and BALA V. BALACHANDRAN
Cost Driver Optimization in Activity-Based Costing **563**

RAJIV D. BANKER and HOLLY H. JOHNSTON
An Empirical Study of Cost Drivers in the U.S. Airline Industry **576**

SRIKANT M. DATAR, SUNDER KEKRE, TRIDAS MUKHOPADHYAY, and KANNAN SRINIVASAN
Simultaneous Estimation of Cost Drivers **602**

NICHOLAS DOPUCH
A Perspective on Cost Drivers **615**

Contents continued on Overleaf

SHORTER ARTICLES

RAY BALL, S. P. KOTHARI, and ROSS L. WATTS
Economic Determinants of the Relation Between Earnings Changes and Stock Returns **622**
MIMI L. ALCIATORE
New Evidence on SFAS No. 69 and the Components of the Change in Reserve Value **639**
WAYNE H. SHAW and HEATHER A. WIER
Organizational Form Choice and the Valuation of Oil and Gas Producers **657**
LAWRENCE D. BROWN and KWON-JUNG KIM
The Association Between Nonearnings Disclosures by Small Firms and Positive Abnormal Returns **668**
MARIO J. MALETTA and THOMAS KIDA
The Effect of Risk Factors on Auditors' Configural Information Processing **681**

BOOK REVIEWS, Shane Moriarity, Editor

Fay, ***Accounting Certification, Educational & Reciprocity Requirements: An International Guide*** .... PAUL E. BAYES **692**
Lehman, ***Accounting's Changing Role in Social Conflict*** .... CRISTI K. LINDBLOM **693**
Rudestan and Newton, ***Surviving Your Dissertation: A Comprehensive Guide to Content and Process*** .... GERALD L. SALAMON **694**
Wallace and Smith, ***Accounting for Options: No Longer an Option*** .... JAMES F. WAEGELEIN **695**
Weetman, editor, ***SSAP 15: Accounting for Deferred Taxation*** .... ROBERT M. KOZUB **695**
Weetman and Gray, editors, ***SSAP 19: Accounting for Investment Properties*** .... JOHN T. RIGSBY **696**
Zeff, Van Der Wel, and Camfferman, ***Company Financial Reporting: A Historical and Comparative Study of the Dutch Regulatory Process*** .. Norlin G. Rueschhoff **697**
List of Books Received **698**
Editorial Policy and Style Information **699**

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1960–1962	Robert K. Mautz
1963–1964	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1989	William R. Kinney, Jr.

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Editor
A. Rashad Abdel-khalik

Associate Editors
Robert H. Ashton
Victor L. Bernard
J. Harry Evans, III
John S. Hughes
Robert P. Magee (*Editor Elect*)
James C. McKeown
Ira Solomon

VOL. 68 OCTOBER 1993 No. 4

ELI AMIR
The Market Valuation of Accounting Information: The Case of Postretirement Benefits other than Pensions **703**
HARRY ZVI DAVIS and YORAM C. PELES
Measuring Equilibrating Forces of Financial Ratios **725**
MARLYS GASCHO LIPE
Analyzing the Variance Investigation Decision: The Effects of Outcomes, Mental Accounting, and Framing **748**
ROBERT M. BUSHMAN and RAFFI J. INDJEJIKIAN
Stewardship Value of "Distorted" Accounting Disclosures **765**
WILLIAM S. WALLER
Auditors' Assessments of Inherent and Control Risk in Field Settings **783**
MARK W. NELSON
The Effects of Error Frequency and Accounting Knowledge on Error Diagnosis in Analytical Review **803**
SANFORD MORTON
Strategic Auditing for Fraud **825**
ELI BARTOV
The Timing of Asset Sales and Earnings Manipulation **840**

SHORTER ARTICLES

ALEX DONTOH and JOSHUA RONEN
Information Content of Accounting Announcements **857**
ROBERT E. VERRECCHIA
How Do We Assess a Model of Price and Volume? **870**
RICHARD C. SANSING
Information Acquisition in a Tax Compliance Game **874**
KWANG-HYUN CHUNG, DIMITRIOS GHICAS, and VICTOR PASTENA
Lenders' Use of Accounting Information in the Oil and Gas Industry **885**
GRACE POWNALL, CHARLES WASLEY, and GREGORY WAYMIRE
The Stock Price Effects of Alternative Types of Management Earnings Forecasts **896**

Contents continued on Overleaf

STEPHEN P. BAGINSKI, EDWARD J. CONRAD, and JOHN M. HASSELL
The Effects of Management Forecast Precision on Equity Pricing and on the Assessment of Earnings Uncertainty **913**
CHARLES R. ENIS
Earnings-Based Compensation Plans, Performance, and Capital Expenditure Policy in the Motor Carrier Industry **928**

COMMENTS

SAMUEL TUNG and JOHN CROWE
Expectation Formation and Financial Ratio Adjustment Processes: A Comment and An Extension **942**
CHENG F. LEE and CHUNCHI WU
Expectation Formation and Financial Ratio Adjustment Processes: A Reply **953**

BOOK REVIEWS, Shane Moriarity, Editor

Aggarwal, editor, ***Capital Budgeting Under Uncertainty*** ..... SUSAN S. HAMLEN **956**
Allen and McDermott, ***Accounting Success: A History of Price Waterhouse in America 1890–1990*** .... R. GLEN BERRYMAN **957**
Anyane-Ntow, editor, ***International Handbook of Accounting Education and Certification*** .... FELIX POMERANZ **958**
Cooper, Kaplan, Maisel, Morrissey, and Oehn, ***Implementing Activity-Based Cost Management: Moving from Analysis to Action*** ... WILLIAM L. FERRARA **959**
Davis and Holt, ***Experimental Economics*** .... RONALD R. KING **960**
Gore, ***The FASB Conceptual Framework 1973–1985: An Analysis*** .... FRANCES L. AYRES **961**
Humphrey, Moizer, and Turley, ***The Audit Expectations Gap in the United Kingdom*** .... STEPHEN TOMCZYK **962**
Ijiri and Nakano, editors, ***Business Behavior and Information*** .... RAMJI BALAKRISHNAN **963**
Neimark, ***The Hidden Dimensions of Annual Reports: Sixty Years of Social Conflict at General Motors*** .... THERESA D. HAMMOND **964**
Pogue, editor, ***State Taxation of Business: Issues and Policy Options*** .... JOHN E. KARAYAN **965**
List of Books Received **966**
Editorial Policy and Style Information **968**
Index to Volume 68 .... **971**

PAST EDITORS

1926–1929	William A. Paton	1968–1970	Charles H. Griffin
1929–1943	Eric L. Kohler	1971–1972	Eldon S. Hendriksen
1944–1947	A. C. Littleton	1973–1975	Thomas F. Keller
1948–1949	Robert L. Dixon	1976–1978	Don T. DeCoster
1950–1959	Frank P. Smith	1978–1983	Stephen A. Zeff
1960–1962	Robert K. Mautz	1983–1987	Gary L. Sundem
1963–1964	Lawrence L. Vance	1987–1989	William R. Kinney, Jr.
1965–1967	Wendell Trumbull		

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 1–27

The Market Response to Pension Plan Terminations

H. Fred Mittelstaedt
University of Notre Dame
Philip R. Regier
Arizona State University

SYNOPSIS AND INTRODUCTION: This study, which analyzes stock returns associated with the announcement to terminate overfunded defined-benefit pension plans, differs from prior research by considering both the role of the replacement plan and prior anticipation of the termination by shareholders. These innovations enable us to (1) evaluate the use of the reversion amount (plan assets minus termination liabilities) as a measure of wealth transfer from plan participants to shareholders, and (2) resolve inconsistencies across prior studies regarding the existence of a significantly positive market reaction at the termination announcement.

This study provides insight into the controversy over the appropriate measure of pension liability: accumulated benefit obligation (ABO) versus projected benefit obligation (PBO). PBO is based on expected future salaries, while ABO is based on current salaries. Barth (1991, 457) suggests that investors consider expected salaries but view PBO as noisy. The current study complements this finding. By examining stock returns at the time of pension plan termination—when any uncertainty regarding the effects of future salaries on pension liability is resolved—we are able to assess whether investors use future salaries to value pension liabilities prior to termination. In addition, the study describes institutional features of pension plans and illustrates the Lanen and Thompson (1988) result that market anticipation may attenuate the relation between the market reaction to and the cash flow effects of an event.

We hypothesize that the change in stock price at the time of termination is directly related to the wealth transfer from plan participants to shareholders. If defined-benefit coverage is continued after termination, the primary wealth transfer is from retired participants because annuity contracts purchased from insurance companies typically do not allow for cost-of-

We would like to acknowledge the helpful comments of the referees, M. Alderson, J. Boatsman, R. Ippolito, S. Kaplan, S. Low, S. Reiter, H. Reneau, M. Stone, and S. Wolcott. We also acknowledge the research assistance of C. Cristea and C. Johnson.

Submitted July 1990.
Accepted July 1992.

living adjustments (COLAs). If defined-benefit coverage is canceled, retired workers lose COLAs, and current workers lose the difference between pension liabilities based on *future* salaries and those based on *current* salaries.

The results are presented in a descriptive, systematic fashion in order to properly associate any market reaction with underlying features of termination and to ascertain the cause of inconsistencies across previous studies. An event-type methodology is employed by using daily CRSP data. Department of Labor Form 5500 data are used in conjunction with disclosures about the nature of the replacement plan to estimate worker losses at the time of termination. In addition, estimates from a censored bivariate probit model are incorporated to control for market anticipation of the termination.

The results suggest that stock returns at the time of termination are affected by the form of the replacement plan and, to a lesser degree, market anticipation. The results are consistent with investors' using future salaries to value pension liabilities prior to termination. Analysis of the replacement plan and comparison of the reversion amount to the theoretically derived wealth transfer suggest that inconsistencies across prior studies principally result from including firms that continue defined-benefit coverage after termination. In addition, the results help explain why previous studies documented small changes in market value relative to the reversion amount.

Key Words: *Pension, Asset reversion, Market anticipation.*

Data Availability: *Subject to licensing agreements, the authors will make the data used in this study available upon request.*

THE remainder of this article is organized as follows. In section I, we discuss institutional features of pension plan terminations and research to date on plan terminations. In section II, an analytical model of the market reaction to plan terminations is provided. We describe the research design in section III and present the results in section IV. Final conclusions and *caveats* are presented in section V.

I. Institutional Features of Pension Plan Terminations and Prior Literature

Institutional Background

Under a defined-benefit pension plan, the employer promises a retirement annuity that is determined according to a formula based on years of service and salary.[1] The employer may also make voluntary cost-of-living adjustments (COLAs) to retirees to counter the effects of inflation.[2] The employer contributes actuarially determined amounts to the plan trust to prefund the pension obligation. Plans can become legally overfunded if actual benefit payments are less than expected or actual returns are higher than expected under actuarial assumptions, or if annual contributions are based on salary levels expected at retirement.

---

[1] The analysis in this article describes a salary-based pension formula. Another type is the flat-benefit formula, which bases pension benefits on years of service times a fixed dollar amount. However, as Ippolito (1991) points out, the fixed dollar amounts are often renegotiated in conjunction with wages at the end of contracts, which effectively indexes the fixed benefits to final salaries. In addition, salary-based formulas are more common than flat-benefit formulas. Turner and Beller (1989, 374) report that of approximately 224,000 plans, 85 percent use salary-based formulas, 8 percent use flat-benefit formulas, and the remaining 7 percent use other or unknown formulas. For 179 terminated plans examined in conjunction with this study, 93 percent used salary-based formulas, 1 percent used flat-benefit formulas, and 6 percent had other or unknown formulas.

In the early 1980s, large numbers of firms began to terminate overfunded defined-benefit pension plans covering currently employed workers. To terminate a pension plan, a firm notifies the Pension Benefit Guaranty Corporation (PBGC) and plan participants of its intent to terminate. Upon termination, firms must satisfy vested and previously unvested liabilities by using salaries in effect at the legal date of termination. The firm then purchases annuities that will satisfy this legal obligation. After government approvals, assets in excess of the legal liabilities (the reversion amount) are returned to the plan sponsor. Typically, defined-benefit coverage is continued or defined-contribution coverage is established for the active participants of the terminated plan. The main terms of the replacement plan are generally disclosed at the time of termination.[3]

Defined-benefit coverage may be continued in one of two ways (see VanDerhei 1987). A *termination/reestablishment* ends the original plan and then establishes a new defined-benefit plan with identical or similar benefits for the affected workers. Because the new plan almost always gives active workers credit for prior service (Hay/Huggins Co., Inc. 1986), the primary difference between the old and new plan is that the new plan has fewer assets. A *spin-off* splits the pension plan into separate parts for active and retired workers. Excess plan assets are moved to the retiree plan, the retiree plan is terminated, plan liabilities are satisfied by purchasing annuity contracts, and the remaining assets revert to the sponsoring firm. Again, the plan for active workers continues as if no termination occurred, except that the plan is less well-funded.

Under either method of continuing defined-benefit coverage, active workers normally receive pension benefits based on anticipated salaries, and retirees are less likely to receive COLAs because their benefits are covered by annuity contracts purchased from insurance companies. Hay/Huggins Co., Inc. (1986, 19–33) reports that for a sample of plans with continuing defined-benefit coverage, less than 10 percent are expected to have lower benefits at retirement. The analysis does not include the effects of lost or reduced COLAs.

If a defined-contribution plan is adopted, it is typically a money purchase plan, a profit-sharing plan, an employee stock ownership plan (ESOP), or a thrift (401K) plan.[4]

---

[2] Allen et al. (1986) report that for a sample of 139,316 U.S. citizens retiring before 1973, 75 percent received at least one COLA between 1973 and 1979. For those receiving COLAs, the mean rate of increase in benefits equaled 51 percent of the increase in the Consumer Price Index.

[3] In 90 percent of the termination announcements appearing in *The Wall Street Journal*, the form of the replacement plan was disclosed. The form of the replacement plan also is typically disclosed in the notification to employees and the PBGC. In addition, according to our discussions with personnel at terminating firms, employees are informed of the proposed replacement plan at the same time they are notified of the termination.

[4] A money purchase plan requires an employer to allocate a fixed portion of an employee's salary to a retirement account. For example, TIAA-CREF is a money purchase plan. Contributions are required regardless of whether the firm is profitable in a particular year. A profit-sharing plan is similar except that a firm has discretion as to the amount and timing of the payments. For instance, in years of losses, the employer may be able to avoid making contributions to the plan. An ESOP may involve either a money purchase plan or a profit

The future cash commitment of the firm differs across plan types and will ultimately depend on the generosity of the plan.

The plan participant may suffer two types of losses when defined-benefit coverage is ended. First, the participant is no longer guaranteed a benefit, and so bears all the investment risk associated with the plan assets. This relieves the firm of some investment risk and possibly reduces the level of fixed payments associated with its pension plan, depending on the type of defined-contribution plan. The second loss to the participant arises because benefits under the terminated pension plan are fixed at the salary levels in effect at the date of termination. Employees are precluded from basing their claims at termination on salaries at retirement. This second type of loss is the primary focus of this study. This type of loss does not arise with defined-benefit plan replacement but can be substantial with defined-contribution plan replacement. Ignoring the effect of voluntary COLAs, Hay/Huggins Co., Inc. (1986, 45–46) estimates that the amount of this loss can exceed 70 percent of the projected benefits that would have been paid under the terminated plan, and 89 percent of the defined-contribution plans are expected to result in lower pension benefits at retirement.

Prior Research on Pension Plan Terminations

One of the main motivations for studying stock returns associated with pension plan terminations is to ascertain whether the termination results in a wealth transfer from employees (Alderson and Chen 1986; Shleifer and Summers 1988). Employees sacrifice wages early in their careers to pay for pension benefits based on *anticipated* salaries (economic liabilities). Upon termination of a plan, firms can meet their legal obligation by paying pension benefits based on *current* salaries (explicit liabilities). The wealth transfer is the difference between the economic and explicit liabilities.[5] The economic liability and explicit liability are analogous to the projected benefit obligation (PBO) and accumulated benefit obligation (ABO), respectively, as defined in *Statement of Financial Accounting Standards No. 87: Employers' Accounting for Pensions* (SFAS No. 87, FASB 1985).[6] This link to SFAS No. 87 and the controversy surrounding the appropriate measure of the pension liability make this study of special interest to accountants.

The fundamental question addressed in previous research—whether positive returns accrue to shareholders in the period surrounding firms' announcements to terminate pension plans—has not been resolved. Such returns would be consistent with investors' valuing pension obligations based on future salary levels. Earlier research generally concluded that positive returns did accrue to sponsoring firms, but later studies noted various specification problems in the earlier works and generally concluded that returns were either much smaller than previously reported or

---

sharing plan. However, to be called an ESOP, the plan must invest at least half of its assets in the common or preferred stock of the sponsoring employer. Finally, a thrift plan calls for voluntary employee contributions or wage deferrals and typically includes some amount of matching portion by the employer. See McGill and Grubbs (1989) for a more complete discussion of the types of defined-contribution plans.

[5] Economic liabilities are often referred to as ongoing or implicit liabilities. Explicit liabilities are often referred to as termination or legal liabilities. To avoid confusion, the terms economic and explicit will be used throughout the remainder of the article.

[6] A key difference between PBO and economic liability is that PBO does not allow voluntary COLAs to be included as part of the obligation. See Warshawsky (1989) for a further discussion of differences between SFAS No. 36 (FASB 1980), SFAS No. 87 (FASB 1985), explicit liabilities, and economic liabilities.

insignificant.[7] These conflicting results remain unresolved, as does the issue of why the change in market value reported in these studies tended to be a small fraction of the average asset reversion (a proxy for the wealth transfer used in the previous studies).

This study addresses these issues by considering factors not incorporated in previous research. First, on the basis of the discussion of differences between defined-benefit and defined-contribution plans, it seems reasonable to expect that stock price reaction will differ depending on the form of the replacement plan. However, none of the previous studies fully considered replacement plan effects.[8] Second, given previous research indicating (1) that management policy is at least partially anticipated by investors (Acharya 1987; Lanen and Thompson 1988; Malatesta and Thompson 1985; Warner et al. 1988), and (2) that terminations can be predicted from financial data (Mittelstaedt 1989; Petersen 1992; Stone 1987; Thomas 1989), it is reasonable to assume partial anticipation by investors of the termination announcement. Such anticipation would reduce the market surprise at the announcement date and attenuate the return response. However, previous studies have neither incorporated anticipation in their theory nor controlled for market expectation in their research design. To our knowledge, this is the first accounting study to apply the Lanen and Thompson (1988) theory by using a bivariate probit model to quantify the degree of market anticipation.

II. Shareholder Wealth Effects Resulting from the Termination Decision

Our model is derived from Ippolito's (1985, 1986a) model of the pension contract and incorporates institutional features discussed in the preceding section.[9] The theory does not center on the issue of whether the firm or the participants own the pension assets, but instead assumes that the firm owns the pension assets; at issue is the nature and amount of the firm's pension obligation. The first part of the analysis outlines two extreme perspectives of the pension contract. To make predictions regarding stock returns at the time of termination, we then consider costs incurred by the firm as a result of the termination and allow for market anticipation.

Two Views of the Pension Contract and Loss at Termination

Let EX and EC represent the two endpoints of a continuum of perspectives regarding the pension obligation. EX represents the explicit perspective which bases the benefits of active participants on *current* salaries and years of service to date. Retiree liabilities are based on current benefit levels and do not assume future COLAs unless these are expressly provided in the pension agreement. Thus, plan participants negotiate salaries under the assumption that management has a right to end the pension plan at any point in time and that COLAs, other than those required by the plan agreement, are entirely voluntary.

---

[7] Alderson and Chen (1986), VanDerhei (1987), Haw et al. (1988) and Hsieh et al. (1990) conclude that relatively positive returns accrue to terminating firms. Conversely, Mitchell and Mulherin (1989), Mittelstaedt and Regier (1990), and Moore and Pruitt (1990) find either insignificant or relatively small returns, and attribute much of the size and significance of the returns to a variety of specification problems in the earlier research.

[8] Exclusion of the form of replacement plan from theory and design affects studies that include terminations after 1983 more than studies using pre-1984 data. Continued defined-benefit coverage became common only after issuance of joint guidelines in May 1984 by the Department of Labor, Internal Revenue Service, and the PBGC. The joint guidelines made it clear that government agencies would approve terminations as long as the guidelines were met. Studies using pre-1984 terminations generally find higher and more significant market reactions than studies using both pre- and post-1984 terminations.

[9] The Ippolito model has been relied upon extensively in pension studies published in economics and finance journals. It also has been referenced in several accounting studies (see, e.g., Francis and Reiter 1987; Mittelstaedt 1989; Stone 1991; Thomas 1989).

In contrast, EC represents the economic perspective under which the obligation for active workers is based on the current level of service and *anticipated* salary levels near retirement. In addition, active employees and retirees expect COLAs throughout retirement even though these are not expressly stated in the plan document (see Ippolito 1986a; Pesando 1984). Thus, plan participants negotiate salaries under the assumption that management *implicitly* agrees to continue the plan indefinitely and to protect retirees from inflation by providing COLAs.

When a pension plan is terminated, the magnitude of the wealth transfer depends on which perspective of the obligation was taken prior to termination and the form of the pension plan replacing the terminated plan. Recall from section I that the firm is legally required to provide only *EX* at termination. If the explicit perspective has been taken prior to termination, no wealth transfer occurs. However, if the economic perspective has been taken, the potential wealth transfer (*PWT*) at termination is:

$$PWT=(EC^{Act}-EX^{Act})+(EC^{Ret}-EX^{Ret}), \tag{1}$$

where *Act* and *Ret* designate active and retired participants, respectively. *PWT* represents the loss to participants from being paid *EX* when *EC* was expected in negotiating salaries. The magnitude of *PWT* is directly related to expected rate of salary increase; the higher the rate, the greater the difference between valuing the pension liability with an economic perspective and valuing the liability with an explicit perspective.

The left-hand side of equation (1) is denoted *PWT* because the actual wealth transfer (*AWT*) depends on the form of the replacement plan. If defined-benefit coverage is ended, then *AWT* equals *PWT* and can be expressed as:

$$AWT^{NDB}=(EC^{Act}-EX^{Act})+(EC^{Ret}-EX^{Ret}), \tag{2}$$

where the superscript *NDB* indicates cessation of defined-benefit coverage. Under our analytical model, no distinction is made for the various types of defined-contribution plans. However, this issue is explored in the empirical portion of the study.

If defined-benefit coverage is continued, active employees still expect pensions based on anticipated salaries because of credit for prior service. However, COLAs are less likely for retired participants because their benefits after termination are provided by annuity contracts purchased from insurance carriers. As a result, the actual wealth transfer arising when defined-benefit coverage is continued can be expressed as:

$$AWT^{DB}=EC^{Ret}-EX^{Ret}, \tag{3}$$

where the superscript *DB* indicates continuation of defined-benefit coverage.

As indicated in section I, prior studies have used the reversion amount to proxy for the wealth transfer. If firms fully fund EC so that plan assets approximate EC at the time of termination, the reversion amount approximates *PWT* and AWT^{NDB}. However, comparison of equations (2) and (3) shows that if defined-benefit plan coverage is continued, the reversion amount overstates AWT^{DB} by the amount $EC^{Act}-EX^{Act}$.

Employee Wealth Transfers and Changes in Stock Prices at the Date of Termination

To predict changes in stock prices at time of termination, costs associated with termination and the market's ability to predict the termination must be considered. Our analysis is derived from the models developed in Lanen and Thompson (1988) and Hand (1991). These models reinforce the intuition that the market's prior anticipation of an event and uncertainty about its costs affect the relation between changes in ultimate cash flows resulting from the event and changes in stock prices.

Prior to termination, the potential wealth transfer can be estimated from financial statements or Form 5500 disclosures.[10] Therefore, for a stock price change to be attributed to the wealth transfer from employees as described above, the investors must view the "true" pension liability as greater than EX. If the market believes that the true liability equals EX, no wealth transfer has taken place and, as a result, no price reaction would take place. For the purpose of modeling, the market's view of the true pension liability is set equal to EC, but the probability of honoring EC is allowed to vary inversely with the costs of not honoring it.

Let Θ^{T} equal the costs of termination when defined-benefit coverage is continued. The costs include legal and administrative fees, immediate vesting, and harm done to firm reputation for not upholding implicit contracts.[11,12] Let Θ^{NDB} equal the incremental costs of discontinuing defined-benefit coverage after termination. Additional costs are expected when defined-benefit coverage is ended because the magnitude of the loss to workers is greater than if it is continued. Assume that at time t_0, managers and investors know that Θ^{T} will take on one of two values at time t_1: Θ^{Tl} and Θ^{Th} with probabilities p_T and $1-p_T$, respectively, where $0 \leq \Theta^{Tl} < EC^{Ret} - EX^{Ret} < \Theta^{Th}$. Also assume they know that if Θ^{Tl} obtains, Θ^{NDB} will take on one of two values, Θ^{NDBl} and Θ^{NDBh}, with probabilities p_{NDB} and $1-p_{NDB}$, respectively, where $0 \leq \Theta^{NDBl} < EC^{Act} - EX^{Act} < \Theta^{NDBh}$. Prior to termination, stock prices reflect the expected value of termination,

$$E(TERMINATE) = [p_T p_{NDB}(EC^{Act} - EX^{Act} - \Theta^{NDBl})] + [p_T(EC^{Ret} - EX^{Ret} - \Theta^{Tl})], \quad (4)$$

where $E(\bullet)$ denotes an expectations operator. The pre-termination probabilities p_T and p_{NDB} are expected to vary with the size of the potential wealth transfer and the financial stress of the firm.[13] The first term in parentheses on the right-hand side of equation (4) is realized only if defined-benefit coverage is ended and it is consequently multiplied by

---

[10] The Employee Retirement Income Security Act of 1974 and the Internal Revenue Code require firms sponsoring qualified asset accumulation plans, namely, defined-benefit and defined-contribution plans, to file Form 5500 annually. The form contains information regarding number of participants, plan income statement, plan balance sheet, actuarial assumptions, etc. The IRS, DOL, and PBGC use the reports to monitor the pension plans of individual firms and to evaluate the U.S. private pension system as a whole. See McGill and Grubbs (1989, 47–75) for a more detailed discussion.

[11] Shleifer and Summers (1988) argue that because many future contingencies are difficult to foresee and measure prior to occurrence, completely explicit contracting to account for every contingency has a prohibitive cost. Accordingly, it is often efficient to form implicit contracts. The practice of allowing employees to receive pensions based on anticipated salaries can be viewed as honoring an implicit contract. Stakeholders of a firm will enter into implicit contracts only if they trust firm management and may become unwilling to agree to future implicit contracts once such contracts are broken, thereby increasing future contracting costs. For this reason, a reputation for upholding implicit contracts can have value.

[12] The primary cost of immediate vesting is that some employees who would have left the firm prior to vesting now must be paid pension benefits. For the full amount of nonvested benefits to be viewed as a cost, the firm must plan to dismiss all nonvested employees prior to vesting or expect all nonvested employees to leave voluntarily prior to vesting. For a sample of 142 terminating plans with complete Form 5500 data (gathered in conjunction with the study), the nonvested benefits represent approximately 7 percent of total explicit benefits ($EX^{Act} + EX^{Ret}$), which represents an upper bound on the cost to the firm. Because nonvested benefits represent a small portion of total benefits and the firm would have difficulty avoiding indefinitely a large portion of these costs in absence of termination, our analysis treats nonvested benefits the same as vested benefits. Separating EX^{Act} into vested and nonvested components reduces slightly the difference between EC^{Act} and EX^{Act}.

[13] As financial stress increases, the long-run viability of the firm and, as a consequence, the value of upholding implicit contracts are questionable. As a result, the costs associated with violating implicit contracts diminishes.

$p_T p_{NDB}$. The second term in parentheses is realized whether defined-benefit coverage is ended or continued after termination and is multiplied by p_T.

Between t_0 and t_1, managers obtain private information that resolves the uncertainty around Θ^T and Θ^{NDB}. With the assumption that managers behave in stockholders' best interests, managers terminate and continue defined-benefit coverage if Θ^{Tl} and Θ^{NDBh} obtain. Managers terminate and end defined-benefit coverage if Θ^{Tl} and Θ^{NDBl} obtain. Stock prices change when costs associated with the termination are revealed to the market by management's announcement to terminate. If defined-benefit coverage is ended,

$$\Delta S^{NDB}=(1-p_T p_{NDB})(EC^{Act}-EX^{Act}-\Theta^{NDBl})+(1-p_T)(EC^{Ret}-EX^{Ret}-\Theta^{Tl}), \tag{5a}$$

where ΔS denotes change in stock price. If defined-benefit coverage is continued,

$$\Delta S^{DB}=-p_T p_{NDB}(EC^{Act}-EX^{Act}-\Theta^{NDBl})+(1-p_T)(EC^{Ret}-EX^{Ret}-\Theta^{Tl}). \tag{5b}$$

If no termination occurs, uncertainty as to whether Θ^{Tl} will obtain in future periods remains unresolved, and stock prices will not change.

Equations (5a) and (5b) form the desired model of the change in shareholder wealth stated as a function of the wealth transfer arising from violation of an implicit contract for active and retired employees, market expectations, the form of the replacement plan, and costs associated with termination. Because $(EC^{Act}-EX^{Act}-\Theta^{NDBl})$ and $(EC^{Ret}-EX^{Ret}-\Theta^{Tl})$ are positive, ΔS^{NDB} must be positive. The sign of ΔS^{DB} depends on the degree of prior anticipation and the relation between the wealth transfer from active employees and retirees. As a result, the first two hypotheses, stated in alternative form, are:

H$_{A1}$: $\Delta S^{NDB}>0$.
H$_{A2}$: $\Delta S^{NDB}>\Delta S^{DB}$.

Although costs of termination are important in modeling the change in stock price, they cannot be measured adequately. If Θ^{NDBl} and Θ^{Tl} are omitted, equations (5a) and (5b) can be rewritten as:

$$\Delta S^{NDB}\leq(1-p_T p_{NDB})(EC^{Act}-EX^{Act})+(1-p_T)(EC^{Ret}-EX^{Ret})=UWT^{NDB}, \tag{6a}$$

$$|\Delta S^{DB}|\leq|-p_T p_{NDB}(EC^{Act}-EX^{Act})+(1-p_T)(EC^{Ret}-EX^{Ret})|=|UWT^{DB}|, \tag{6b}$$

where UWT is the unexpected wealth transfer from plan participants to shareholders before transaction costs. Both sides of equation (6b) are expressed as absolute values because UWT^{DB} may be positive or negative, depending on the relation between the active and retiree portions. Change in stock price is expected to be related directly to the unexpected wealth transfer that arises when the firm pays EX instead of EC. However, because costs cannot be measured, the relation is less than one-to-one.

From this discussion, the third hypothesis can be stated as:

H$_{A3}$: $|\Delta S|\leq|UWT|$,

where ΔS represents the change in stock price for firms continuing or ending defined-benefit coverage.

III. Empirical Procedures

Sample Selection

The primary source for termination announcements in this study is the PBGC notification tape. The Dow Jones News Service (DJNS) and *The Wall Street Journal*

Index are used to locate public announcements of terminations. In the initial sample, 158 firms on the PBGC notification tape met the following criteria:

1. The notification concerns a verifiable termination of a primary plan, and is not associated with asset sales, plant closings, discontinued operations of the sponsoring company, or a recently acquired subsidiary;
2. The reversion amount is in excess of $1 million; and
3. The PBGC notification date falls between October 1, 1981 and December 31, 1987.

The first criterion isolates the effects of the termination event. As Thomas (1989) indicates, terminations tied to a control change may be different from those not associated with such a change. The second and third criteria result from data limitations.

From the initial sample of 158 firms reporting terminations, we excluded those with a notification or publication date that exceeded the mandatory notification of employees by more than four days (48 firms). For three firms, the type of replacement plan could not be identified. Nine firms were missing Form 5500 information, ten firms were missing COMPUSTAT data, and six firms had insufficient CRSP data. The final sample consists of 82 firms that terminated pension plans between 1982 and 1987. Forty-one of the firms continued defined-benefit coverage, and 41 firms ended this coverage. There were five terminations in 1982, 13 in 1983, 23 in 1984, 26 in 1985, 13 in 1986, and two in 1987.

Change in Stock Price

Measure of Returns. All the hypotheses require estimation of ΔS^{NDB} and ΔS^{DB}. These variables are operationalized by procedures discussed in Patell (1976). For each security j, the market model is estimated and an abnormal return for day t is computed as:

$$AR_{jt} = R_{jt} - (a_j + b_j R_{mt}), \tag{7}$$

where R_{jt} is the return on security j, R_{mt} is the value-weighted market return, and a_j and b_j are OLS coefficient estimates from regressing j's individual daily returns on value-weighted market returns over 300 days (days -360 to -61). Day 0 is the termination announcement date. Cumulative abnormal returns (CAR_t) are calculated by summing abnormal returns across four cumulation intervals indicated by the subscript t: $t=1$ for announcement date returns, 5 for returns summed from day -2 to 2, 10 for returns summed from day -4 to 5, and 15 for returns summed from day -4 to 10. These intervals allow for event date misspecification and the possibility that it may take investors several days to estimate the effects of the replacement plan (the range of event dates is discussed further below). The average CAR_t for companies ending defined-benefit coverage is denoted $CAR_{NDB,t}$; the same variable for companies continuing defined-benefit coverage is $CAR_{DB,t}$.

Selection of Event Date. Following Mitchell and Mulherin (1989) and Mittelstaedt and Regier (1990), we use the earlier of either the public announcement or the PBGC notification as the event date. Prior to the passage of the Single-Employer Pension Plan Amendments Act of 1986 (SEPPAA), employees *and* the PBGC had to be notified at least ten days prior to the legal date of termination, the date at which liabilities become fixed. According to personnel at the PBGC, the notification was stamped as received, and it is assumed to coincide with the mandatory notification to employees. Thus, this date would reasonably represent the first time the market could have learned of the termination and the replacement plan absent prior disclosure (a public announcement in

The Wall Street Journal or on the DJNS). When the public announcement precedes the notification date, there is a high probability that this is the date of initial disclosure. However, using only firms with publication dates in *The Wall Street Journal* or on the DJNS limits the sample size because not all firms publicly announce terminations.[14] As a result of these considerations, the date used reflects the trade-off between the precision of event date specification and sample size.

After the passage of SEPPAA, employees had to be notified 60 days prior to the legal date of termination, but the PBGC could be notified "as soon as practicable" after employees are notified, so PBGC notification may no longer coincide with mandatory notification of employees. When the PBGC notification follows the mandatory employee notification date, the PBGC notification is a poor event-date, and we exclude terminations with a notification date or public announcement more than four days after the mandatory employee notification date. We also begin the event interval at day −4 for the 10-day and 15-day event periods.

Confounding Events. Many termination announcements are made with or near other important firm announcements, such as restructuring of operations, restructuring of debt, or concerns about takeover (Mittelstaedt and Regier 1990). The issue of confounding events is not as critical in this study as in previous studies because our chief concern is the association between wealth transfer proxies and returns, not simply whether terminations result in positive returns. When the 82-firm sample is reduced by as much as 40 percent to eliminate firms with confounding news events, the association between wealth transfer proxies and returns is similar to the results for the full sample. With the reduced sample, the null form of hypothesis H_{A1} is rejected, but the null form of hypothesis H_{A2} is not rejected.

Independent Variables Used to Test Hypothesis H_{A3}

Hypothesis H_{A3} addresses the relation between the unexpected wealth transfer at termination and stock prices. The components of the unexpected wealth transfer are EC^{Act}, EX^{Act}, EC^{Ret}, EX^{Ret}, p_T, and $p_T p_{NDB}$. The economic and explicit liabilities for active and retired workers are operationalized by transforming plan information reported in Form 5500. This information includes the assumed interest rate for calculations of pension obligation, the market value of the trust assets, and the accrued benefit liability. Although the Form 5500 liability is based on current salaries, firms have discretion in selecting the assumed interest rate. As a result, to obtain a measure of explicit liability, reported liability must be adjusted by using an interest rate that would be obtained if retirement annuities were purchased. The Form 5500 data were used instead of SFAS No. 36 or SFAS No. 87 data because the former are measured at the plan level, and separate disclosure is provided for retired and active participants. Economic liabilities are estimated by using a real rate of interest instead of the nominal rate used to compute the Form 5500 liability.

We use the Ippolito method to adjust reported liabilities to economic or explicit measures. Ippolito (1986a) derives a relation between pension liabilities and other characteristics of pension plans reported on Form 5500 such as assumed interest rates, number of actives and retirees, and annual retirement benefits, then uses nonlinear least squares and Form 5500 data from 1978 to estimate parameters for the following expression:

---

[14] For the 82 firms in the final sample, only 39 percent made public announcements. Of these, 78 percent of the news dates precede the PBGC notification date.

$$\frac{L_{5500}}{C_{5500}P_{5500}}=\beta_1 e^{-\beta_2 I_{5500}^{Act}}+\beta_3\frac{N_{5500}}{C_{5500}}e^{-\beta_4 I_{5500}^{Ret}}+\epsilon, \tag{8}$$

where L_{5500} is the total reported pension liability, C_{5500} is the reported number of active participants, P_{5500} is the reported average annual pension, N_{5500} is the reported number of retired participants, I_{5500}^{Act} is the reported assumed interest rate for active liabilities, and I_{5500}^{Ret} is the reported assumed interest rate for retiree liabilities.[15] The coefficients β_1 and β_3 are functions of average age of active workers, average age of retirees, average expected date of retirement, and average expected length of retirement. The coefficients β_2 and β_4 are needed to obtain discount factors for the liabilities for active employees and retirees, respectively. The estimated coefficients for β_2 and β_4 are 0.077 and 0.057, respectively.[16] Although the higher coefficient for active workers is consistent with pension payments being farther in the future for active employees than those for retirees, Ippolito reports that the difference between the coefficients is not significant.

With the Ippolito parameters, reported liabilities can be converted to economic and explicit liabilities. To convert reported liabilities to economic liabilities:

$$EC^{Act}=L_{5500}^{Act}\times e^{[-.077\times(r-I_{5500}^{Act})]}; \tag{9}$$

$$EC^{Ret}=L_{5500}^{Ret}\times e^{\{-.057\times[(r+.5i)-I_{5500}^{Ret}]\}}. \tag{10}$$

To convert reported liabilities to explicit liabilities:

$$EX^{Act}=L_{5500}^{Act}\times e^{[-.077\times(Aaa-I_{5500}^{Act})]}; \tag{11}$$

$$EX^{Ret}=L_{5500}^{Ret}\times e^{[-.057\times(Aaa-I_{5500}^{Ret})]}, \tag{12}$$

where L_{5500}^{Act} and L_{5500}^{Ret} are the reported liabilities for active and retired participants, respectively; r is the real interest rate; *Aaa* is the average Moody's rate for the year of termination; and *i* is the expected inflation rate, calculated as the difference between *Aaa* and r.[17] The Ippolito conversion method has been used in several studies to estimate economic and explicit pension obligations (Francis and Reiter 1987; Ippolito 1985, 1986b; Ippolito and James 1992; Reiter 1991).[18]

---

[15] Interest rates enter the regression and subsequent adjustment process as percentages; for example, a rate of 8 percent enters the equation as 8, not 0.08.

[16] It is possible that the cross-sectional relation estimated by Ippolito with 1978 data is not stable during the experimental period of the study. To assess the stability of the relation, we reestimated Ippolito's model using 4,075 observations from Form 5500 for 1984, the midpoint of the experimental period of this study. The estimates for β_2 and β_4 when using 1984 data are 0.079 and 0.050, respectively. The similarity of our results to Ippolito's results suggests that the relation between interest rates and pension liabilities remained stable from 1978 to 1984.

[17] It should be noted that the calculation of economic liabilities assumes that salaries increase at the rate of inflation. Barth (1991, 455) reports that 42 percent of 1,042 firms use salary progression assumptions that are higher or lower than expected inflation.

[18] A second method uses the formula that plan sponsors are required by the PBGC to calculate to determine the variable portion of the company's insurance premium. The objective of the PBGC in applying the formula is to adjust for cross-firm differences in actuarial rate assumptions. However, it also has been used in pension research to adjust reported pension liabilities to a common interest rate (Bodie and Papke 1990; Petersen 1992). Liability estimates with the Ippolito method are less sensitive to choice of interest rate than those derived from the PBGC method, but the values obtained from both procedures over a range of interest rates have correlation coefficients approaching 1 (greater then 94 percent). Because the results are similar with either method, results obtained with the PBGC method are not reported.

To simplify the exposition, the results presented in the next section reflect use of an assumed real interest rate of 1.5 percent. However, we examined the sensitivity of the results to interest rate assumptions at various stages in the empirical work. The primary regression results are qualitatively similar to those reported regardless of method used when the real interest rate is varied between extreme values of 1 and 5 percent.[19] For example, increasing the real interest rate from 1 to 5 percent reduces the size of the actual wealth transfer (see eqs. [2] and [3] above) as a percentage of total market value of the firm from 8.3 percent to 6.3 percent; it does not, however, change substantially the size of significance of any of the regression coefficients reported in table 4.

Mittelstaedt (1989), Petersen (1992), Stone (1987), and Thomas (1989) suggest that the termination decision is positively related to degree of financial stress. Stone (1991) concludes that the decision to end defined-benefit plan coverage (hereafter called the *NDB* decision) also is affected by financial stress.

The findings of these prior studies are used to develop probit models for the purpose of estimating investors' evaluation of p_T and $p_T p_{NDB}$.[20] New models are estimated because previous models typically use data prior to 1985, and none could generate $p_T p_{NDB}$. The new models are estimated with samples that are larger than those used previously and include terminations from 1981 to 1988. In addition, two innovations are made to estimate $p_T p_{NDB}$. First, the *NDB* prediction model includes a dummy variable to account for the 1984 change in regulation that first permitted firms to continue defined-benefit plans after termination. Second, a bivariate probit model is utilized to acknowledge that defined-benefit coverage can be ended *only after* a termination.

In conformance with the econometric work of Meng and Schmidt (1985) and Boyes et al. (1989), the termination and *NDB* decisions can be modeled as:

$$y_1 = X_1\beta_1 + e_1, \quad y_1 = \begin{cases} 1 \textit{ if plan is terminated,} \\ 0 \textit{ if plan is not terminated,} \end{cases} \tag{12}$$

$$y_2 = X_2\beta_2 + e_2, \quad y_2 = \begin{cases} 1 \textit{ if defined-benefit coverage is ended,} \\ 0 \textit{ if defined-benefit coverage is continued,} \end{cases} \tag{13}$$

where X_1 and X_2 are matrices of independent variables for the termination and *NDB* prediction models, respectively. Use of the bivariate model offers efficiency gains over separate estimation of equations (12) and (13) and allows for the potential correlation, ρ, between the two equations. Consideration of the correlation corrects for the potential sample selection bias that could arise from estimating equation (13) separately (Boyes et al. 1989, 6). P_T equals $\phi(X_1\beta_1)$, and $p_T p_{NDB}$ equals $F(X_1\beta_1, X_2\beta_2, \rho)$, where β_1, β_2, and ρ

---

[19] During the early 1980s, real interest rates were at historically high levels; the ten-year moving average at the end of 1986 was 4.3 percent when calculated as the difference between Moody's average *Aaa* bond rate and the percentage change in the consumer price index. Ippolito discounts the vested obligations for active employees at 1 percent and retiree benefits at 1.5 percent plus 50 percent of expected inflation (Ippolito 1986a, 66). The justification for not indexing retiree benefits fully to inflation is based on the experience of the 1970s, when firms increased retiree benefits by only about 50 percent of inflation (Allen et al. 1986). Equation (10) in the text follows this procedure.

[20] Discussions of the factors influencing the termination and *NDB* decisions are found in previously published works. Therefore, the discussion here deals primarily with operationalizing the theoretical constructs identified in prior work. In many cases, the variables are defined in the same way as in other studies. The main purpose for developing the models in this study is to estimate prior probabilities, not to test theories regarding the factors influencing the decision to terminate.

are estimated by using the log-likelihood function described in Meng and Schmidt (1985), and $F(\bullet)$ and $\phi(\bullet)$ denote the bivariate standard normal c.d.f. and univariate c.d.f., respectively.

The dependent variables for the termination and *NDB* prediction models are obtained by examining *PBGC* documents and firms' financial statements. If a firm terminates a pension plan, y_1 is coded 1; otherwise, it is coded 0. If a firm terminates a plan and ends defined-benefit coverage, y_2 is coded 1; otherwise, it is coded 0. We classified a firm as a terminator according to criteria similar to those described in the first part of this section.

The independent variables for the termination and *NDB* prediction models are measured as of year −1, with year 0 defined as the termination year for terminators and a randomly assigned year for nonterminators. The random assignment is based on the calendar-year distribution observed for terminators.

The termination prediction model includes tax, cash flow, dividend, debt, overfunding, and change in tax law variables. A tax status dummy variable is coded 0 if a firm has unused net operating tax loss carryforwards, receives a refund for tax loss carrybacks, or does not pay taxes; otherwise, the variable is coded 1. Cash flow is defined as cash flows from operations divided by total assets. The dividends variable is defined as cash dividends paid to common equity scaled by total assets. The debt variable is defined as the market value of common stock divided by debt. Overfunding is defined as the difference between pension assets and explicit liabilities divided by the market value of common stock. Explicit liabilities are estimated from the aggregate pension liabilities reported under SFAS No. 36 by using the Ippolito (1986a) approach described earlier. Because the specific numbers of retirees and active employees are not reported in SFAS No. 36 disclosures, we use the Francis and Reiter (1987) assumptions that 75 percent are active employees and 25 percent are retirees. All of these independent variables are computed with COMPUSTAT data. The final variable, tax law change dummy, is coded 1 if year 0 is after 1985; otherwise, it is coded 0. The variable takes into account the introduction in 1986 of a ten percent excise tax on the reversion amount.

Petersen (1990) studies pension plans that have not been terminated and finds that the probability of sponsoring a defined-benefit plan is positively related to the number of employees and negatively related to downside variability of income. As a consequence, the *NDB* prediction model includes the number of employees, downside variability of income, and change in pension regulations. Number of employees is defined as the natural log of the number of firm employees (in thousands). Downside variability is defined as the number of times in a five-year period that income is less than half of the five-year average. The pension regulation change dummy is coded 1 if year 0 is prior to 1984; otherwise, it is coded 0. The *NDB* prediction model includes a set of variables that is different from the one used in the termination model to ensure identification of the equations.

Active and research COMPUSTAT tapes are used to identify firms in the primary, supplementary, tertiary, and OTC files with SFAS No. 36 asset information for 1981 to ensure that all sample firms sponsor defined-benefit plans. To acknowledge the fact that firms can obtain a reversion only if they have overfunded plans, firms that are underfunded in the aggregate are eliminated (results are similar when underfunded plans are included in the analysis). The final sample of overfunded firms includes 865

Table 1
Bivariate Termination and *NDB* Prediction Models

	Coefficient	$\lvert t \rvert$	*Two-tailed Probability*
Termination Prediction Model:			
Intercept	−0.989	8.650	0.000
Tax dummy	−0.221	1.829	0.067
Cash flow/assets	−0.272	0.454	0.650
Dividends/assets	−3.659	0.812	0.417
Market value/debt	−0.081	2.865	0.004
Overfunding/market value	1.867	5.270	0.000
Tax law change dummy	0.021	0.168	0.867
NDB Prediction Model:			
Intercept	0.074	0.105	0.916
Number of employees	−0.277	2.821	0.005
Downside variability of income	−0.077	0.559	0.557
Pension regulation change dummy	1.281	3.246	0.001
p-value	0.269	0.785	0.432

nonterminators and 132 terminators. Seventy-three of the 132 terminators end defined-benefit coverage, and 59 continue it. To avoid biased estimates caused by choice-based sampling (see Zmijewski 1984), all nonterminators meeting the data requirements are used to estimate the models.

We present the model parameters for the termination and *NDB* prediction models in table 1. Chi-squared tests indicate that the models are highly significant. In conformance with prior studies, the decision to terminate is negatively related to tax status and the ratio of market value to debt and is positively related to the amount of overfunding relative to market value. The cash flow and dividend variables are insignificant, which may be due to the correlation with other variables in the model. The tax law change dummy also is insignificant.

The decision to end defined-benefit coverage is negatively related to the number of employees and positively related to a dummy indicating whether the event year was before 1984. Although Petersen (1990) finds downside variability to be positively related to sponsorship of defined-contribution plans, the variable is insignificant in the *NDB* decision. The correlation betweeen equations is also insignificant. As a result, the model parameters change very little when the equations are estimated independently instead of jointly. However, this result was not known *a priori*.

The reversion amount is included in several analyses to evaluate its ability as a proxy for the wealth transfer from employees at termination. The reversion amount is obtained from the *PBGC* notification tape.

IV. Results

In this section we provide statistical summary measures and analyses of stock return and wealth transfer variables, including (1) descriptive data regarding the wealth

Table 2

Mean and Standard Deviation for Variables of Liability, Wealth Transfer, Probability, and Returns

Variables	*Firms Continuing Defined-Benefit Coverage (DB)*	*Firms Ending Defined-Benefit Coverage (NDB)*	p-*values*
Liability (Form 5500 data)			
Active employees	0.101 (0.084)	0.155 (0.256)	0.212
Retirees	0.090 (0.095)	0.096 (0.187)	0.860
Net asset reversion	0.124 (0.098)	0.179 (0.257)	0.207
Potential wealth transfer (*PWT*)	0.102 (0.085)	0.150 (0.229)	0.214
Active employees	0.086 (0.073)	0.134 (0.205)	0.168
Retirees	0.016 (0.017)	0.016 (0.030)	0.929
Probability of terminating (p_T)	0.184 (0.092)	0.221 (0.187)	0.254
Probability of terminating and ending *DB* ($p_T p_{NDB}$)	0.077 (0.039)	0.147 (0.147)	0.005
Unexpected wealth transfer (*UWT*)	0.005 (0.014)	0.101 (0.127)	<0.001
Active employees			
$-p_T p_{NDB}(EC^{Act}-EX^{Act})$	−0.007 (0.008)		<0.001
$(1-p_T p_{NDB})(EC^{Act}-EX^{Act})$		0.092 (0.117)	
Retirees			
$(1-p_T)(EC^{Ret}-EX^{Ret})$	0.012 (0.014)	0.009 (0.014)	0.265
Returns			
$CAR_{DB(NDB),1}$	0.001 (0.021)	0.015** (0.030)	0.009
$CAR_{DB(NDB),5}$	0.000 (0.049)	0.016* (0.049)	0.070
$CAR_{DB(NDB),10}$	−0.014 (0.093)	0.016* (0.066)	0.047
$CAR_{DB(NDB),15}$	−0.005 (0.109)	0.025* (0.101)	0.097

Notes: Means are followed by standard deviations in parentheses. The variables of liability, reversion, and wealth transfer are scaled by total market value of equity five days prior to the announcement to terminate. The liability figures taken directly from Form 5500 do not reflect adjustment to common interest rates. The p-values indicate the statistical probability of error in rejecting the null hypothesis of no difference between the values for *DB* and *NDB* firms. Two-tailed tests were used for variables of liability, reversion, and wealth transfer; one-tailed tests of *NDB* greater than *DB* were used for variables of probability and return.

* Significantly greater than zero at 0.10 alpha level.

** Significantly greater than zero at 0.01 alpha level.

transfer variables and empirical tests of hypotheses H_{A1} and H_{A2}, (2) step-by-step testing of hypothesis H_{A3} with regressions of stock returns on alternative measures of wealth transfer, and (3) a series of specification checks.

Descriptive Statistics and Tests of Hypotheses H_{A1} and H_{A2}

Summary statistics for estimated variables and statistical results related to the first two hypotheses are given in table 2. Results are partitioned on the basis of whether firms continue or end defined-benefit coverage. As shown, liability, reversion, and *PWT* amounts are a large percentage of total market value whether a firm continues or ends defined-benefit coverage. The mean values of these variables are similar across the two classes of firms. When *PWT* is decomposed, the larger amount for active employees reflects both the large number of active employees relative to retirees, and the greater spread between *EC* and *EX* for active employees because of the difference between current and anticipated salaries.

The results also demonstrate the empirical relation between the reversion amount, *PWT*, and *AWT*. The magnitude of the asset reversion and *PWT* are similar. When parametric and nonparametric matched-pairs tests are conducted (not shown), the null hypothesis of no difference cannot be rejected for either group of terminators. Thus, the reversion amount appears to be a reasonable proxy for *PWT*. Because *PWT* equals *AWT* for firms ending defined-benefit coverage (see eq. [2]), the reversion amount is also a reasonable proxy for *AWT* for firms ending defined-benefit coverage. However, for firms continuing defined-benefit coverage, *AWT* equals *PWT* from retirees (see eq. [3]). For firms continuing defined-benefit coverage in our sample, *PWT* is approximately six times higher than *AWT* (0.102/0.016). The reversion amount overstates *AWT* by a factor of approximately eight (0.124/0.016) for these firms. *AWT* for firms continuing defined-benefit coverage is approximately 11 percent (0.016/0.150) of *AWT* for firms ending defined-benefit coverage.

Summary statistics for p_T and $p_T p_{NDB}$ for the 82-firm sample indicate differences in probabilities between firms that continue and those that end defined-benefit coverage, which means that the procedures used to estimate the probabilities appropriately discriminate between the two types of firms. For firms ending defined-benefit coverage, p_T is not significantly different from that of firms continuing defined-benefit coverage, but $p_T p_{NDB}$ is significantly higher (*p*-value of 0.005).[21] It is assumed that investors would have formed similar probabilities prior to termination.

Adjusting *AWT* for differences in prior probabilities in accordance with equations (6a) and (6b) reduces the measure of wealth transfer from 0.016 to 0.005 for firms continuing defined-benefit coverage and from 0.150 to 0.101 for firms ending defined-benefit coverage. This indicates that market anticipation further reduces the ability of the reversion amount to proxy for the wealth transfer at termination, regardless of whether firms end or continue defined-benefit coverage.

The results for returns in table 2 support rejection of the null form of hypothesis H_{A1}. Depending on the length of cumulation, mean excess returns for the firms ending defined-benefit coverage range from 1.5 to 2.5 percent, and are greater than zero at a ten percent alpha level across all event intervals. This evidence is consistent with the conclusion that reversions result in positive returns for firms that eliminate defined-benefit coverage. Similarly, the p-values support rejection of hypothesis H_{A2} by indicating that the returns for firms ending defined-benefit coverage are greater than those of firms continuing such coverage. The one-tailed probabilities for the 1-day and 10-day cumulation periods are less than 0.05, and the probabilities for the 5-day and 15-day periods are less than 0.10.

The results suggest that there is a positive market reaction to termination when defined-benefit coverage is ended but not when it is continued. Although early studies of termination overstated the degree of market reaction because of design weaknesses, their conclusion of a significant positive reaction was correct, but only for a subset of

---

[21] For the sample of 865 nonterminating firms used to estimate the termination and *NDB* prediction models, p_T and $p_T p_{NDB}$ are 0.12 and 0.068, respectively.

Table 3
Mean and Standard Deviation for Daily Return Variables

	Firms Continuing Defined-Benefit Coverage (DB)	Firms Ending Defined-Benefit Coverage (NDB)	t-statistic p-values	Wilcoxon p-values
$CAR_{DB[NDB],Day\ -4}$	−0.008 (0.026)	−0.003 (0.027)	0.207	0.432
$CAR_{DB[NDB],Day\ -3}$	0.001 (0.018)	0.003 (0.025)	0.361	0.192
$CAR_{DB[NDB],Day\ -2}$	−0.005 (0.017)	−0.003 (0.027)	0.324	0.658
$CAR_{DB[NDB],Day\ -1}$	−0.001 (0.018)	−0.008 (0.027)	0.908	0.822
$CAR_{DB[NDB],Day\ 0}$	0.001 (0.021)	0.015** (0.030)	0.009	0.005
$CAR_{DB[NDB],Day\ 1}$	−0.001 (0.017)	0.002 (0.032)	0.318	0.655
$CAR_{DB[NDB],Day\ 2}$	0.006* (0.020)	0.010** (0.024)	0.206	0.189
$CAR_{DB[NDB],Day\ 3}$	−0.001 (0.019)	−0.006 (0.022)	0.846	0.944
$CAR_{DB[NDB],Day\ 4}$	−0.005 (0.023)	0.004* (0.017)	0.034	0.023
$CAR_{DB[NDB],Day\ 5}$	−0.001 (0.018)	0.003 (0.019)	0.160	0.158
$CAR_{DB[NDB],Day\ 6}$	−0.001 (0.019)	0.007* (0.027)	0.060	0.066
$CAR_{DB[NDB],Day\ 7}$	0.009* (0.030)	0.001 (0.026)	0.891	0.624
$CAR_{DB[NDB],Day\ 8}$	−0.002 (0.021)	−0.002 (0.033)	0.491	0.452
$CAR_{DB[NDB],Day\ 9}$	−0.002 (0.018)	0.004 (0.028)	0.157	0.219
$CAR_{DB[NDB],Day\ 10}$	0.004 (0.035)	−0.001 (0.028)	0.797	0.696

Note: Means are followed by standard deviations in parentheses. The *t*-test and Wilcoxon p-values indicate one-tailed probabilities of error in rejecting a hypothesis of *DB* greater than *NDB* values.

* Significantly greater than zero at 0.10 alpha level.
** Significantly greater than zero at 0.01 alpha level.

terminating firms. More recent studies corrected many of the earlier design problems, but did not consider the role of replacement plans. As shown in table 2, including firms that continue defined-benefit coverage will bias the findings toward insignificance. The results also are consistent with the market's considering future salaries in valuing pension liabilities prior to termination. Positive returns result only when the effect of salary growth on the size of the pension liability is removed.

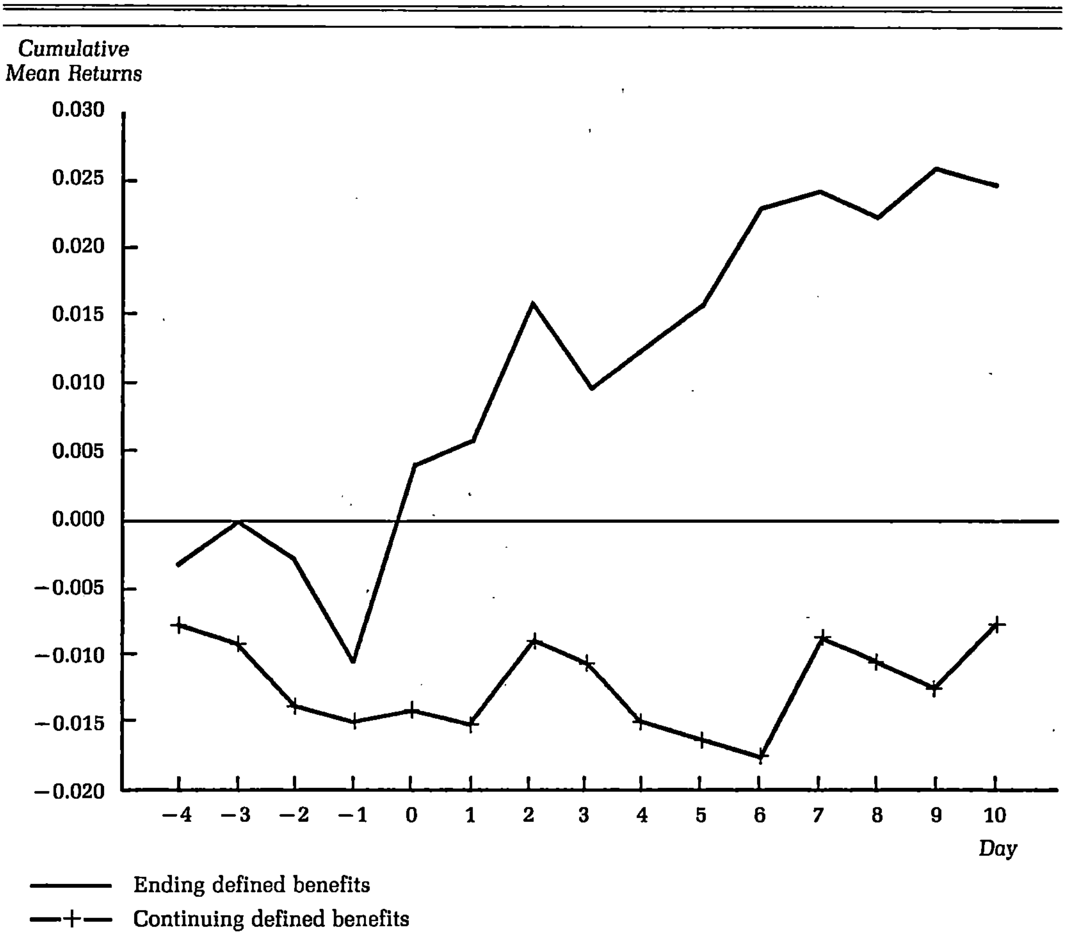
The return results are most significant on day 0. Given the difficulty in isolating the announcement date, this result requires greater explication. In table 3, we report the daily mean returns within the 15-day period, and show that firms ending defined-benefit coverage had statistically positive returns on four of the 15 days. Figure 1 graphs these mean returns, cumulated from day −4, and reinforces the finding that returns were greater throughout the 15-day period for terminating firms ending defined-benefit coverage than for terminating firms continuing such coverage. The Wilcoxon results reported in table 3 confirm that the significant mean differences in returns on days 0, 4, and 6 across firms continuing and ending defined-benefit coverage are not the result of extreme observations.

Tests of Hypothesis H_{A3}

Hypothesis H_{A3} predicts a relation between returns and the unexpected wealth transfer at termination. To show the effects of using increasingly refined measures of the wealth transfer, we report results for the reversion amount (*REVERS*), *PWT*, *AWT*, and *UWT* in table 4.

Figure 1

Plot of Mean Abnormal Returns Cumulated from Day −4



For the all-firms category, R^2 values and coefficients on the independent variable generally increase in moving down the table, which is consistent with increased explanatory power of the wealth transfer proxy. For the full sample, there appears to be little advantage to using *PWT* rather then *REVERS* to explain returns, but there is a marked increase in the size and significance of the coefficient and R^2 values once knowledge of the replacement plan is incorporated (*AWT*). These results are consistent with the market's viewing *REVERS* and *PWT* as upward-biased measures of the actual wealth transfer at termination. As noted for table 2, the overstatement arises principally from firms continuing defined-benefit coverage. Including the effects of market anticipation (*UWT*) increases the size and significance of the wealth transfer coefficients and marginally improves R^2 values. Consistent with hypothesis H_{A3}, the coefficient relating *UWT* to returns is significantly greater than 0 and less than 1.

Comparison of *REVERS* and *PWT* results for the firms ending defined-benefit coverage indicates that *PWT* is more highly correlated than *REVERS* with changes in

Table 4

Cross-Sectional Univariate Regressions of Abnormal Returns on the Net Reversion, Potential Wealth Transfer, Actual Wealth Transfer, and Unexpected Wealth Transfer

Independent Variable	Interval (t)	*All Firms* ($n=82$)			*Firms Ending Defined Benefits* ($n=41$)			*Firms Continuing Defined Benefits* ($n=41$)		
		a	b	R^2	a_{NDB}	b_{NDB}	R^2	a_{DB}	b_{DB}	R^2
Reversion amount (REVERS)		$\tilde{CAR}_{k,t}=a+b\,\tilde{REVERS}_k+\tilde{\epsilon}_k$			$\tilde{CAR}_{i,t}=a_{NDB}+b_{NDB}\tilde{REVERS}_i+\tilde{\epsilon}_i$			$\tilde{CAR}_{j,t}=a_{DB}+b_{DB}\tilde{REVERS}_j+\tilde{\epsilon}_j$		
	1	0.000 (0.993)	0.051 (<0.001)	0.14	0.005 (0.355)	0.056 (0.001)	0.24	0.003 (0.636)	−0.014 (0.653)	<0.01
	5	−0.000 (0.994)	0.054 (0.026)	0.05	0.006 (0.495)	0.056 (0.031)	0.09	−0.000 (0.972)	0.004 (0.477)	<0.01
	10	−0.007 (0.555)	0.052 (0.132)	0.02	0.004 (0.773)	0.071 (0.040)	0.08	0.006 (0.814)	−0.158 (0.851)	0.03
	15	−0.001 (0.966)	0.069 (0.126)	0.02	0.005 (0.779)	0.110 (0.038)	0.08	0.031 (0.262)	−0.290 (0.951)	0.07
Potential wealth transfer (*PWT*)		$\tilde{CAR}_{k,t}=a+b\,\tilde{PWT}_k+\tilde{\epsilon}_k$			$\tilde{CAR}_{i,t}=a_{NDB}+b_{NDB}\tilde{PWT}_i+\tilde{\epsilon}_i$			$\tilde{CAR}_{j,t}=a_{DB}+b_{DB}\tilde{PWT}_j+\tilde{\epsilon}_j$		
	1	0.002 (0.665)	0.050 (0.001)	0.11	0.006 (0.219)	0.056 (0.002)	0.19	0.004 (0.443)	−0.031 (0.781)	0.02
	5	−0.000 (0.991)	0.065 (0.019)	0.05	0.005 (0.532)	0.071 (0.016)	0.11	0.002 (0.850)	−0.021 (0.591)	<0.01
	10	−0.009 (0.418)	0.081 (0.062)	0.03	−0.000 (0.991)	0.109 (0.007)	0.14	0.008 (0.724)	−0.216 (0.893)	0.04
	15	−0.007 (0.602)	0.137 (0.021)	0.05	−0.003 (0.874)	0.185 (0.003)	0.18	0.024 (0.359)	−0.291 (0.925)	0.05

(*Continued*)

Table 4—*Continued*

Cross-Sectional Univariate Regressions of Abnormal Returns on the Net Reversion, Potential Wealth Transfer, Actual Wealth Transfer, and Unexpected Wealth Transfer

Independent Variable	*Interval* (t)	*All Firms* ($n=82$)			*Firms Ending Defined Benefits* ($n=41$)			*Firms Continuing Defined Benefits* ($n=41$)		
		a	b	R^2	a_{NDB}	b_{NDB}	R^2	a_{DB}	b_{DB}	R^2
Actual wealth transfer (AWT)		$\tilde{CAR}_{k,t}=a+b\tilde{AWT}_k+\tilde{\epsilon}_k$			$\tilde{CAR}_{i,t}=a_{NDB}+b_{NDB}\tilde{AWT}_i+\tilde{\epsilon}_i$			$\tilde{CAR}_{j,t}=a_{DB}+b_{DB}\tilde{AWT}_j+\tilde{\epsilon}_j$		
	1	0.003 (0.372)	0.062 (<0.001)	0.17	0.006 (0.219)	0.056 (0.002)	0.19	0.004 (0.407)	−0.189 (0.822)	0.02
	5	0.002 (0.783)	0.079 (0.005)	0.08	0.005 (0.532)	0.071 (0.016)	0.11	−0.002 (0.783)	0.196 (0.338)	<0.01
	10	−0.009 (0.358)	0.122 (0.009)	0.07	−0.000 (0.991)	0.109 (0.007)	0.14	0.000 (0.980)	−0.925 (0.851)	0.03
	15	−0.005 (0.669)	0.183 (0.003)	0.09	−0.003 (0.874)	0.185 (0.003)	0.18	0.021 (0.367)	−1.671 (0.948)	0.07
Unexpected wealth transfer (UWT)		$\tilde{CAR}_{k,t}=a+b\tilde{UWT}_k+\tilde{\epsilon}_k$			$\tilde{CAR}_{i,t}=a_{NDB}+b_{NDB}\tilde{UWT}_i+\tilde{\epsilon}_i$			$\tilde{CAR}_{j,t}=a_{DB}+b_{DB}\tilde{UWT}_j+\tilde{\epsilon}_j$		
	1	0.003 (0.399)	0.098 (0.001)	0.14	0.006 (0.290)	0.086 (0.009)	0.14	0.001 (0.702)	−0.094 (0.651)	<0.01
	5	0.000 (0.995)	0.153 (0.002)	0.10	0.002 (0.819)	0.139 (0.010)	0.13	−0.005 (0.562)	0.923 (0.046)	0.07
	10	−0.012 (0.237)	0.240 (0.003)	0.09	−0.006 (0.655)	0.216 (0.004)	0.17	−0.016 (0.300)	0.476 (0.327)	0.01
	15	−0.009 (0.477)	0.352 (0.001)	0.12	−0.013 (0.491)	0.372 (0.001)	0.22	−0.003 (0.857)	−0.388 (0.623)	<0.01

Notes: All variables are scaled by the market value of equity prior to the termination announcement. The coefficients are followed by p-values in parentheses; these are two-tailed for intercepts, one-tailed for other coefficients. There are four cumulation intervals; $t=1$ for announcement day returns, 5 for returns summed from day −2 to 2, 10 for returns summed from day −4 to 5, and 15 for returns summed from day −4 to 10.

market expectations; R^2 values and wealth transfer coefficients generally increase. This result is consistent with the wealth effect's being tied to a change in obligation rather than to the amount of assets that reverted to the firm. The effects of controlling for market anticipation are similar to those observed for the all-firms category.

The results in table 4 do not show *AWT* or *UWT* to be superior to *REVERS* in explanatory power for the firms continuing defined-benefit coverage, possibly because of the small magnitude of the wealth transfer for these firms. From table 2, *AWT* is estimated to be only 1.6 percent of market value. When market anticipation is considered, the estimated wealth transfer (*UWT*) drops to 0.5 percent. When this amount is reduced by costs associated with termination, the net result may be so small that no link between returns and *AWT* (or *UWT*) can be demonstrated with the procedures used in this study.

It is important to note that although the link between *AWT* and stock returns is not supported for the firms continuing defined-benefit coverage, using *AWT* or *UWT* instead of *REVERS* or *PWT* for these firms increases the size of the coefficient for the all-firms sample. This result and other results in table 4 suggest that the form of the replacement plan and market anticipation are important in explaining the market response to termination.

Relative Explanatory Power of the Four Proxies

An important result implied but not directly shown in table 4 is that the actual and unexpected wealth transfer proxies explain more variation in returns than the reversion amount used in prior studies. This result was directly tested in the following way. First, three difference variables were calculated: The reversion amount minus the potential wealth transfer (*REVERS−PWT*), the potential wealth transfer minus the actual wealth transfer (*PWT−AWT*), and the actual wealth transfer minus the unexpected wealth transfer (*AWT−UWT*). These variables were then used as explanatory variables in three bivariate regressions of returns on the reversion amount and *REVERS−PWT*, the potential wealth transfer and *PWT−AWT*, and the actual wealth transfer and *AWT−UWT*. The significance of these difference variables is a measure of the additional explanatory power provided by the more refined proxy (Kmenta 1986, 596). Because there was no *a priori* relation between the reversion amount and the potential wealth transfer, the sign of *REVERS−PWT* is not predicted. However, the potential wealth transfer is a less refined measure of the actual wealth transfer, which is in turn a less refined measure of the unexpected wealth transfer. Accordingly, the coefficients on *PWT−AWT* and *AWT−UWT* are predicted to be negative.

Results of these regressions are provided in table 5. As suggested by examination of coefficient and R^2 values in table 4, the potential wealth transfer does not provide incremental explanatory power vis-à-vis the reversion amount for the 1-, 5-, and 10-day intervals and provides only marginal improvement in the 15-day period. Consideration of the replacement plan incorporated in the actual wealth transfer provides significant improvement relative to the potential wealth transfer across all intervals. Finally, incorporating expectations provides a marginal improvement in explanatory power for the 5-, 10-, and 15-day returns; no improvement is provided for the day 0 returns.

The results in tables 4 and 5 are consistent with the market's reassessing the effect of future salaries on pension obligations in that there is a direct relation between our

P6241

Table 5

Cross-Sectional Bivariate Regressions of Abnormal Returns on Wealth Transfer and Differences Between Wealth Transfer

Interval (t)	Regression 1 $[\tilde{CAR}_{k,t}=a+b\tilde{REVERS}_k +c(\tilde{REVERS}_k-\tilde{PWT}_k)+\tilde{\epsilon}_k]$				Regression 2 $[\tilde{CAR}_{k,t}=a+b\tilde{PWT}_k +c(\tilde{PWT}_k-\tilde{AWT}_k)+\tilde{\epsilon}_k]$				Regression 3 $[\tilde{CAR}_{k,t}=a+b\tilde{AWT}_k +c(\tilde{AWT}_k-\tilde{UWT}_k)+\tilde{\epsilon}_k]$			
	a	b	c	R^2	a	b	c	R^2	a	b	c	R^2
1	−0.000 (0.978)	0.054 (0.001)	−0.008 (0.768)	0.14	0.005 (0.141)	0.057 (<0.001)	−0.111 (0.004)	0.18	0.003 (0.348)	0.049 (0.112)	0.028 (0.632)	0.17
5	−0.001 (0.908)	0.067 (0.018)	−0.047 (0.392)	0.06	0.005 (0.520)	0.074 (0.010)	−0.134 (0.050)	0.09	−0.000 (0.981)	0.184 (0.010)	−0.233 (0.074)	0.10
10	−0.008 (0.471)	0.079 (0.068)	−0.098 (0.282)	0.03	0.003 (0.818)	0.103 (0.024)	−0.336 (0.006)	0.10	−0.012 (0.226)	0.303 (0.011)	−0.403 (0.066)	0.09
15	−0.004 (0.776)	0.129 (0.028)	−0.221 (0.058)	0.06	0.007 (0.664)	0.163 (0.007)	−0.400 (0.010)	0.12	−0.009 (0.461)	0.419 (0.006)	−0.527 (0.061)	0.12

Notes: All variables are scaled by the market value of equity prior to the termination announcement. Coefficients are followed by *p*-values in parentheses; these are two-tailed for intercepts and *REVERS*−*PWT*, one-tailed for other coefficients. There are four cumulation intervals; $t=1$ for announcement day returns, 5 for returns summed from day −2 to 2, 10 for returns summed from day −4 to 5, and 15 for returns summed from day −4 to 10.

estimate of the wealth transfer and abnormal stock returns. In addition, the results reinforce our prior findings that the reversion amount is not a good proxy for the wealth transfer from plan participants when defined-benefit coverage is continued. This is one reason for the less than one-to-one relation between the reversion and abnormal returns documented in prior studies. The results also suggest that consideration of market expectations improves the link between the wealth transfer proxy and abnormal returns, but not to the same degree as consideration of the form of the replacement plan.

Additional Results and Specification Issues

This section examines (1) the incremental informativeness provided by dividing the wealth transfer at termination between active and retired workers, (2) the possibility that the results are attributable to factors associated with a shift from defined-benefit coverage other than the difference between economic and explicit benefits, and (3) relative measurement error and the ability of four empirical measures to proxy for the wealth transfer at termination.

Separate Coefficients for Active Employees and Retirees. Although the preceding results establish the usefulness of assessing the form of the replacement plan for active workers in explaining return variation associated with terminations, other data are available to make a finer assessment of what causes the wealth transfer at termination. The results in table 2 suggest that the observed wealth transfer arising when firms replace defined-benefit plans is driven principally by a wealth transfer from active employees, not retirees. To examine this issue, *UWT* was split into its active and retiree components and the models in table 4 were reestimated. The results (not shown) indicate that the significant association between the wealth transfer and returns observed in table 4 for firms ending defined-benefit plans is indeed driven primarily by the component of the wealth transfer from active employees. The coefficient for this transfer is highly significant, but the coefficient for the retiree component is never significant. However, as mentioned earlier, the lack of significance could be because of the small size of the retiree component.

Alternative Explanation for Results. Our theory suggests that the increase in stock returns is directly related to the difference between pension benefits based on anticipated salaries and benefits based on current salaries. But the firm also may reduce risk by ending defined-benefit plans as discussed in section I. Thus, the reaction could be attributed to the decline in risk. To investigate this issue, the type of defined-contribution plan is incorporated in two separate analyses.

First, mean *CAR*s were compared for firms which established (1) money purchase plans and all other plans, and (2) ESOPs, thrifts, or no plan and all other plans. Money purchase plans tend to shift the least risk to plan participants, while ESOPs, thrifts, and no plan tend to shift the most risk to plan participants.[22] The exact effects of the plan types cannot be known without very detailed information about the subsequent plan and a time-series of contributions to the replacement plans. There were no statistically significant differences in *CAR*s across the four cumulation periods examined, indicating that the results for hypotheses H_{A1} and H_{A2} are not driven by a shift in risk, as measured by type of defined-contribution plan.

---

[22] For the 41 firms ending defined-benefit coverage, 14 established money purchase plans, 11 established profit sharing plans, six established ESOPs, four established thrift plans, four established defined-contribution plans that could not be identified by type, and two firms did not establish replacement plans.

Second, cross-sectional regressions of the forms contained in table 4 were estimated for firms that ended defined-benefit coverage, with the addition of an explanatory indicator variable representing the switch to either (1) an ESOP, thrift, or no plan; or (2) a money purchase plan. These coefficients were not significant, which suggests that the market's primary concern is with whether firms end defined-benefit plans; the form of the subsequent defined-contribution plan does not appear to be informative.

In addition, the intercept terms in the additional analyses and the regressions reported in table 4 are not significantly different from 0. If there is a material shift in risk that is uncorrelated with *UWT*, it seems that the intercept should be positive for the firms ending defined-benefit coverage. It is possible that the benefit associated with a reduction in investment risk could be correlated with *UWT*. However, we believe that the market is reacting to the worker loss arising from the firm's satisfying economic liabilities by paying explicit liabilities and that shift in investment risk is a secondary effect. This view is also consistent with labor's concern that the amount of retirement benefits are lower under subsequent plans (U.S. Congress 1988, 181–87).

Relative Measurement Error of Wealth Transfer Proxies. Our theory predicts that the coefficient on *UWT* is less than 1 because of costs associated with termination.[23] However, measurement error also may be lowering the coefficient. To establish upper bounds for *UWT* and obtain insights into the relative measurement error in the wealth transfer proxies, reverse regression is employed.

In reverse regression, the original explanatory variable (in this case, the wealth transfer proxy) is regressed on the original dependent variable (in this case, *CAR*), thereby placing the measurement error of the wealth transfer proxy into the disturbance term (Beaver et al. 1987). The inverse of the coefficient for *CAR* obtained in the reverse regression represents an upper bound for the coefficients on the wealth transfer proxies estimated from forward regression. One of the properties of reverse regression is that the inverse of the reverse regression coefficient equals the forward regression coefficient divided by the R^2 obtained in the forward regression. Thus, the upper bound for the *UWT* coefficient in the regression for the all-firms sample, 1-day interval, is 0.70 (0.098/0.14).

Bounds were computed for the regressions for all firms and firms ending defined-benefit plans. The results (not shown) suggest that measurement error may be contributing to the size of the *b* coefficients. For the 1-day interval, the upper bound for *UWT* is 0.70 for the regression for all firms, and 0.61 for the regression for firms ending defined-benefit plans. When the event interval is increased, the upper bounds exceed 1 for both samples. This finding suggests that the low coefficients are affected by measurement error as well as by costs associated with termination.

The range between the upper and lower bounds provides some indication of the seriousness of the measurement problem (Kmenta 1986, 352). We compared the ranges of the *REVERS*, *AWT*, and *UWT* coefficients (the *PWT* was not included because of its similarity to *REVERS* and because it equals *AWT* for firms that end defined-benefit coverage). For the full sample, *AWT* has the narrowest range and *UWT* has a wider range, which suggests that consideration of the replacement plan reduces measurement error, but inclusion of market anticipation (at least as implemented in our study) increases measurement error. This result also is consistent with the finding shown in table 5 that *AWT* consistently improves explanatory power, whereas improvement in moving from *AWT* to *UWT* is marginal. The results for the firms that end defined-benefit coverage are similar to those for the full sample except that the ranges for *REVERS* and *AWT* are similar. However, this is expected, given that *AWT* equals *PWT* for these firms; further, prior results suggest that *PWT* and *REVERS* are substitutes.

---

[23] By assuming Θ^{TI} and Θ^{NDBI} are equal across subsamples, they can be estimated by inserting the means of variables from table 2 into equations (5a) and (5b) and solving them simultaneously. For the four cumulation periods, the estimates of Θ^{TI} are 2 to 3 percent of stock value, and estimates of Θ^{NDBI} are 10 to 11 percent.

V. Conclusion

Previous research in the area of pension terminations viewed the amount and ownership of reverted assets as paramount in explaining equity returns at termination. This study adopts the different perspective that the nature and amount of pension liabilities are critical in understanding returns at termination.

In discussing the nature of the pension liability, we provide a theoretically defensible set of wealth transfer variables that function as alternatives to the reversion amount used in previous studies to explain returns. We hypothesize the importance of the replacement plan in formulating the wealth transfer function, and empirically demonstrate that returns at termination are linked to the replacement plan. Much of the inconsistency in results across previous studies may be explained through consideration of the replacement plan and the inadequacy of the reversion amount as an appropriate proxy for the wealth transfer. Additionally, we attempt to control for the attenuating effects of market anticipation by estimating prior probabilities of termination and discontinuation of defined-benefit coverage and incorporating them in the research design. The results are also consistent with investors' valuing the effects of anticipated salaries on pension plan obligations prior to termination.

Principal limitations of our study include the inability to proxy costs in the empirical analysis and control for other factors associated with a shift from defined-benefit coverage. The presence of costs associated with termination is suggested by the fact that not all legally overfunded plans eligible for termination are terminated. Not incorporating such costs causes the empirically observed relation between returns and the unanticipated wealth transfer to be understated. In addition, it is possible that factors associated with a shift from defined-benefit coverage other than those identified in our theory may have affected results.

References

Acharya, S. 1987. Insider's decision of corporate events: A simple measure of expected abnormal return. Working paper, New York University.

Alderson, M., and K. Chen. 1986. Excess asset reversions and shareholder wealth. *Journal of Finance* 41 (March): 225–41.

Allen, S., R. Clark, and D. Sumner. 1986. Postretirement adjustments of pension benefits. *The Journal of Human Resources* 21 (Winter): 118–37.

Barth, M. 1991. Relative measurement errors among alternative pension asset and measures. *The Accounting Review* 66 (July): 433–63.

Beaver, W., R. Lambert, and S. Ryan. 1987. The information content of security prices: A second look. *Journal of Accounting & Economics* 9 (July): 139–57.

Bodie, Z., and L. Papke. 1990. *Pension Portfolio and Investment Strategies: Final Report for Department of Labor Contract Number J-9-P-8-0097.*

Boyes, W., D. Hoffman, and S. Low. 1989. An econometric analysis of the bank credit scoring problem. *Journal of Econometrics* 40 (January): 3–14.

Financial Accounting Standards Board. 1980. *Statement of Financial Accounting Standard No. 36: Disclosure of Pension Information*. Stamford, CT: FASB.

———. 1985. *Statement of Financial Accounting Standard No. 87: Employers' Accounting for Pensions*. Stamford, CT: FASB.

Francis, J., and S. Reiter. 1987. Determinants of corporate pension funding strategy. *Journal of Accounting & Economics* 9 (April): 35–60.

Hand, J. 1991. Not LIFO again?!—A theoretical and empirical re-examination of 1974–1975 LIFO adoptions and non-adoptions. Working paper, University of Chicago.

Haw, I., W. Ruland, and A. Hamdallah. 1988. Investor evaluation of overfunded pension plan terminations. *Journal of Financial Research* 11 (Spring): 81–88.

Hay/Huggins Co., Inc. 1986. *Pension Plan Terminations With Asset Reversions: Final Report for Department of Labor Contract Number J-9-P-4-0092*.

Hsieh, S., K. Ferris, and A. Chen. 1990. Securities market response to pension fund termination. *Contemporary Accounting Research* 6 (Spring): 550–72.

Ippolito, R. 1985. The labor contract and true economic pension liabilities. *American Economic Review* 75 (December): 1031–43.

———. 1986a. *Pensions, Economics, and Public Policy*. Homewood, IL: Dow Jones-Irwin.

———. 1986b. The economic burden of corporate pension liabilities. *Financial Analysts Journal* 42 (January-February): 22–34.

———. 1991. Encouraging long-term tenure: Wage tilt or pensions? *Industrial and Labor Relations Review* 44 (April): 520–35.

———, and W. James. 1992. LBOs, reversions, and implicit contracts. *The Journal of Finance* 47 (March): 139–67.

Kmenta, J. 1986. *Elements of Econometrics*. New York: Macmillan.

Lanen, W., and R. Thompson. 1988. Stock price reactions as surrogates for the net cash flow effects of corporate policy decisions. *Journal of Accounting & Economics* 10 (December): 311–34.

Malatesta, P., and R. Thompson. 1985. Partially anticipated events: A model of stock price reactions with an application to corporate acquisitions. *Journal of Financial Economics* 14 (June): 237–50.

Meng, C., and P. Schmidt. 1985. On the cost of partial observability in the bivariate probit model. *International Economic Review* 26 (February): 71–85.

McGill, D. M., and D. S. Grubbs. 1989. *Fundamentals of Private Pensions*. 6th ed. Homewood, IL: Irwin.

Mitchell, M., and J. Mulherin. 1989. The stock price response to pension terminations and the relation of terminations with corporate takeovers. *Journal of Financial Management* 18 (Autumn): 41–56.

Mittelstaedt, H. F. 1989. An empirical analysis of the factors underlying the decision to remove excess assets from overfunded pension plans. *Journal of Accounting & Economics* 11 (November): 399–418.

———, and P. R. Regier. 1990. Further evidence on excess asset reversions and shareholder wealth. *Journal of Risk and Insurance* 62 (September): 471–86.

Moore, N. H., and S. W. Pruitt. 1990. A comment on excess asset reversions and shareholder wealth. *The Journal of Finance* 45 (December): 1709–14.

Patell, J. 1976. Corporate forecasts of earnings per share and stock price behavior: Empirical tests. *Journal of Accounting Research* 14 (Autumn): 246–76.

Pesando, J. 1984. Employee evaluation of pension claims and the impact of indexing initiatives. *Economic Inquiry* 22 (January): 1–17.

Petersen, M. 1990. Financial aspects of a firm pension choice: DB or not DB. Working paper, University of Chicago.

———. 1992. Pension reversions and worker-stockholder wealth transfers. *Quarterly Journal of Economics* 107 (August): 1033–56.

Reiter, S. 1991. Economic measures of unfunded pension obligations. *Quarterly Review of Economics and Finance*.

Shleifer, A., and L. Summers. 1988. Breach of trust in hostile takeovers. In *Corporate Takeovers: Causes and Consequences*, edited by A. Auerbach. Chicago: University of Chicago Press.

Stone, M. 1987. A financing explanation for overfunded pension plan terminations. *Journal of Accounting Research* 25 (Autumn): 317–26.

———. 1991. Firm financial stress and pension plan continuation/replacement decisions. *Journal of Accounting and Public Policy* 10 (Fall): 175–206.

Thomas, J. 1989. Why do firms terminate overfunded pension plans? *Journal of Accounting & Economics* 11 (November): 361–98.

Turner, J., and D. Beller. 1989. Appendix I—Tables of selected aspects of pensions. In *Trends in Pensions*, edited by J. Turner and D. Beller. Washington, D.C.: U.S. Department of Labor.

U.S. Congress. Senate. Joint Committee on Overfunding and Underfunding of Pension Plans. 100th Cong., 1st sess., 1988. S. Hrg. 100–431.

VanDerhei, J. 1987. The effect of voluntary termination of overfunded pension plans on shareholder wealth. *Journal of Risk and Insurance* 59 (March): 131–56.

Warner, J., R. Watts, and K. Wruck. 1988. Stock prices and top management changes. *Journal of Financial Economics* 20 (March): 461–92.

Warshawsky, M. 1989. The adequacy of funding of private defined benefit pension plans. In *Trends in Pensions*, edited by J. Turner and D. Beller. Washington, D.C.: U.S. Department of Labor.

Zmijewski, M. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22 (Supplement): 59–82.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 28–47

Complementarity of Prior Accounting Information: The Case of Stock Dividend Announcements

Rajiv D. Banker
University of Minnesota
Somnath Das
University of California at Berkeley
Srikant M. Datar
Stanford University

SYNOPSIS AND INTRODUCTION: We present empirical evidence that prior accounting information, such as capital expenditure, retained earnings, funds from operations, and dividend history, is useful in explaining cross-sectional variations in the market response to stock dividend announcements. An important accounting issue concerns the information content of disclosures and their usefulness to the investor. We demonstrate the complementary role of previously disclosed firm-specific accounting information in the market's assessment of subsequently disclosed information. Thus, two firms declaring the same amount of stock dividend may experience predictably different market reactions to the announcement when it is conditioned by prior information about the firms.

Comparable research by Kane et al. (1984) has shown that changes in earnings and dividends either corroborate or contradict prior information. A broader approach by Gonedes (1978) and Antle et al. (1991) shows that the sequence and history of information arrival are relevant in interpreting the information content of accounting signals. Ou and Penman (1989) demonstrate the role of prior accounting information in predicting earnings changes in subsequent periods, and John and Lang (1991) have shown, both theoretically and empirically, that the market uses information about

This article is based in part on Somnath Das's doctoral dissertation work at Carnegie Mellon University. An earlier version was circulated under the title "Differential Information Content of Stock Dividend Announcements." The comments and suggestions of seminar participants at the University of Alberta, University of Arizona, Boston University, University of California at Berkeley, Carnegie Mellon University, Columbia University, Dartmouth College, University of Florida, New York University, Ohio State University, Purdue University, Washington University (St. Louis), and University of Wisconsin–Madison, and detailed comments by William Beaver, Yuji Ijiri, Baruch Lev, Mark Mazur, and anonymous referees are gratefully acknowledged.

Submitted February 1990.
Accepted June 1992.

prior insider trading to interpret the information content of dividend changes.

Stock dividends are appropriate for an investigation of the complementary role of prior accounting information because their issuance is largely a paper transaction, and because they have been interpreted as a signal of better future prospects. Although significant positive abnormal returns usually accompany stock dividend announcements, alternative (and more credible) instruments could signal future prospects (such as an increase in cash dividends). The uncertainty about how investors interpret stock dividend distributions suggests a role for previously disclosed accounting information as a conditioning factor.

A survey of managers of firms declaring stock dividends (Eisemann and Moses 1978) indicates that such distributions are intended either to *conserve cash* in difficult times or to *express confidence* in the firm, two diametrically opposed motivations. So one firm may declare stock rather than cash dividends in order to invest in more profitable ventures, and another may do so because it faces operating losses and a severe cash crunch.[1]

Absent other information, it is likely that the market will respond negatively when cash dividends are discontinued or decreased and replaced by stock dividends (Shefrin and Statman 1984). Other firm-specific accounting information, however, such as capital expenditure (which reflects new investments) and funds from operations (which reflect cash availability), may also influence investor responses when considered in conjunction with dividend history.

Key Words: *Complementary information, Stock dividends, Information content, Market response.*

Data Availability: *Data used in the study are available in the quarterly and annual industrial COMPUSTAT files and the Center for Research in Security Prices (CRSP) daily and master files. A list of company names can be obtained from the authors.*

THE remainder of this article is organized in four sections. First, we develop hypotheses regarding investor response to stock dividend announcements and the use of previously disclosed accounting information. Second, we describe the data and the sample used to test the hypotheses. We discuss the research methodology and empirical results in the third section, and concluding remarks are presented in the fourth section.

[1] An example of the former category in Lakonishok and Lev (1987, 923) is that of Braniff Airlines, which declared a 3 percent stock dividend in 1971 "to conserve cash for investment in new airplanes." An illustration of the latter is taken from a news story in *The Wall Street Journal* (21 October 1981, 38): "The Board of Directors of Wickes Companies, Inc. today declared a 2.5 percent stock dividend payable March 10,1982. . . . This is in lieu of the quarterly cash dividend, which has been 26 cents per share . . ." Related news items indicated that Wickes was experiencing operating losses and cash flow problems.

I. Development of Hypotheses

In developing testable hypotheses regarding the differential market response to the announcement of a stock dividend, we employ residual returns from the one-factor market model as a measure of the response, in conformance with most prior studies of information content.

Prior Dividend History

To the extent that stock dividend announcements change investors' expectations of a firm's future cash dividends, a change in its market valuation will occur.[2] The direction and size of the change will depend on how the market forms its expectations. Our basic thesis is that the market will respond differentially (in both sign and magnitude of the abnormal return) to an announcement, depending on whether the market interprets the firm as (1) having significant growth opportunities or (2) being short on cash and substituting stock dividends for cash dividends. To make this distinction, the market uses other information already released as contradictory or corroborative evidence. Our main concern is with the differential intensity of the market's reaction to stock dividend announcements rather than with the reaction, *per se*, positive or negative.[3]

Aharony and Swary (1980) and Dielman and Oppenheimer (1984) document a negative market reaction to cash dividend omissions and decreases. That is, the market interprets decreases in cash dividends as bad news, rather than as an attempt to conserve cash to fund future growth. A decrease or discontinuation of cash dividends is sometimes accompanied by an initiation of stock dividends, and we interpret these conditions as bad news, following Shefrin and Statman (1984), who argue that investors prefer cash dividends.

For a firm with no prior cash dividend history or one that merely maintains its cash dividend levels, a decreased level of stock dividend is also bad news since it implies a retraction of confidence-building signals about future prospects. However, absent other information, a stock dividend announcement that maintains or increases the past level and that is preceded by maintenance or increase of cash dividends is likely to be interpreted as good news. We classify prior dividend history as "bad" if cash dividends are decreased some time prior to the stock dividend announcement.[4] If cash dividends are maintained immediately prior to the stock dividend announcement *and* if stock dividends are increased or maintained, then prior dividend history is classified as "good." Table 1 displays the classification scheme and the frequency of sample observations in each category.[5]

---

[2] This is consistent with the results reported by Fama et al. (1969, 20), who examined a sample of stock splits. Their "evidence suggests that in reacting to a split, the market reacts only to its dividend implications. That is, the split causes price adjustments only to the extent that it is associated with changes in the anticipated level of future dividends."

[3] In a similar vein, one could argue that firms in financial difficulty use stock dividend announcements to differentiate themselves favorably from other firms in worse difficulty and unable to issue stock dividends. Stock dividend announcements by firms short on cash could signal expectations that the current financial difficulty will be overcome in the future.

[4] For our empirical classification, this time length varies from one to three years, depending on the frequency of a firm's dividend announcements.

[5] Alternatively, firms can be classified by good or bad history according to whether the cumulative abnormal return in a 30-day period prior to the announcement is positive or negative. The basic results and conclusions remain unchanged when such a classification is used.

Table 1
Classification of Firms by Cash and Stock Dividend History

	Announced Stock Dividend			
Cash Dividend History	*Decreased*	*Maintained*	*Increased*	*No Prior Dividend*
Discontinued or decreased	Bad history (1)	Bad history (9)	Bad history (0)	Bad history (11)
Maintained	Bad history (5)	Good history (41)	Good history (0)	Good history (9)
Increased	Bad history (0)	Good history (0)	Good history (0)	Good history (0)
No prior dividend	Bad history (12)	Good history (72)	Good history (0)	Good history (27)

Note: The numbers in parentheses correspond to the number of sample observations.

We use the cash and stock dividend history to identify firms according to whether we expect a strong or weak response to stock dividend announcements. Conditioned by the cash and stock dividend history, the magnitude of the stock market reaction is then explained by using other previously disclosed but complementary sources of firm-specific accounting information, such as the level of capital expenditures and fund flows.

Size of Stock Distribution

In a signaling equilibrium framework, firm managers choose the level of distribution such that the marginal benefit (i.e., marginal increase in market value) exceeds the marginal (signaling) cost associated with the distribution. Therefore, one would expect the magnitude of abnormal returns to increase with the size of the distribution. In particular, Leland and Pyle (1977), Kalay (1980), and Talmor (1981) have shown that size of cash dividend distributions varies positively with the firm's true value. If these theoretical results based on cash dividends are extended to stock dividends, then the magnitude of the distribution is expected to affect the magnitude of abnormal return positively in a signaling equilibrium framework.[6] Grinblatt et al. (1984) use the size of the distribution as an explanatory variable in their examination of abnormal returns around announcement dates and find a positive, but statistically insignificant, association for stock dividends. Woolridge (1983) and McNichols and Dravid (1990) also find a positive relation between increases in mean abnormal return and the size of the stock distribution.

[6] McNichols and Dravid (1990) argue that a Spence-type signaling equilibrium can explain average market reaction to stock dividend announcements. The nature of the signaling cost for stock dividends, other than restriction of "free reserves," is not well-known. Sosnick (1961), however, documents that transaction costs associated with stock dividends increases with the size of the distribution.

The above argument suggests that the size of the distribution is an important determinant of the magnitude of positive abnormal return for firms with a good history. However, for firms with a bad history, larger stock dividends are likely to impose greater costs on the firm, and hence the market reaction is likely to be increasingly negative with the size of the distribution.[7]

The above analysis suggests the following hypothesis:

H1: The anouncement period abnormal return is positive (negative) for firms with a good (bad) history. The absolute magnitude of the announcement period abnormal returns is increasing in the size of the stock dividend.

Retained Earnings Constraint

Stock dividends receive a different accounting treatment than cash dividends and stock splits. Stock dividends and splits are distributions of common stock that increase the number of shares outstanding. In this sense, stock dividends and splits do not directly affect the firm's cash flow, unlike cash dividends. Stock dividends, however, have accounting implications that do not apply to a stock split. According to generally accepted accounting principles, firms declaring stock dividends are required to transfer an amount equal to the market value of the distributed shares from retained earnings to stockholders' equity (contributed capital) in their financial accounting statements.[8] This transfer may impose costs on the firm.[9] For example, creditors often impose constraints on the level of retained earnings to be maintained by a firm. Thus, a firm with a large "free reserve" would find it less costly to declare stock dividends than one with a small or no reserve.

Several previous researchers[10] have argued that firms expecting increases in earnings are likely to declare stock dividends, whereas those not expecting increases will find it costly to do so. The argument is that the special accounting treatment of stock dividends constrains the ability of management to distribute cash dividends in the future if the firm is subject to debt covenants. In earlier research to test this hypothesis, the variable of interest is the level of retained earnings *per se* and not the level of unrestricted retained earnings. However, it is well-known that debt covenants and level of treasury shares often restrict the level of retained earnings that a firm has to maintain. Most commonly, a dividend pool called "unrestricted retained earnings" is explicitly defined in debt agreements. Since a stock dividend entails a transfer from retained earnings, what is relevant is the level of unrestricted retained earnings and how that will be affected by a distribution. The relative reduction in unrestricted retained earnings as a result of a stock dividend declaration measures the extent to which an existing restriction will become more binding. A detailed discussion of the measurement of this variable is provided in the appendix.

---

[7] For example, even though the contemporaneous market reaction to a stock dividend announcement is negative, the declaration may be interpreted as favorable complementary information in evaluating a subsequent financial disclosure. Also note that the modeling of incentives for a large stock dividend by firms with a bad history is beyond the scope of this study.

[8] Accounting Research Bulletin No. 43 is the authoritative document that dictates the appropriate accounting treatment. The guideline for distinguishing between stock splits and stock dividends provided by the AICPA (1961) is that distributions up to 20 through 25 percent should ordinarily be regarded as stock dividends and those above this as splits. No attempt is made here to test the validity of the AICPA rule.

[9] Previous research has generally referred to this as the "retained earnings hypothesis." In this study, we simply assume that the transfer rule in conjunction with bond covenants makes for a credible signal.

[10] See Grinblatt et al. (1984), Lakonishok and Lev (1987), Asquith et al. (1989), and McNichols and Dravid (1990).

It is possible that the retained earnings constraint affects firms in different ways. For firms with a good history, the signaling argument indicates that the expected positive market reaction will increase with the relative reduction in unrestricted retained earnings. Large relative reductions in unrestricted retained earnings are costly, and only firms with very favorable prospects are likely to make such large declarations. For firms with a bad history, however, the market will not interpret the stock dividend declaration as a signal of better future prospects if the retained earnings constraint appears more binding. In fact, the announcement may result in a negative reaction, since greater restrictions are being put on the level of free reserves.

Stated as a testable hypothesis:

H2: The announcement period abnormal return is increasing with the relative reduction in the level of unrestricted retained earnings for firms with a good history, decreasing for firms with a bad history.

Level of Capital Expenditure

In their financial accounting textbook, Davidson et al. (1988, 523) state, "A stock dividend formalizes the fact that some of the funds represented by past earnings have been used for plant expansion, to replace assets at increased prices, or to retire bonds. Such funds are therefore unavailable for cash dividends." This suggests that the relative level of capital expenditure may indicate whether a firm is indeed capitalizing its already-invested funds as permanent capital stock. Firms with relatively high levels of capital expenditure may have a more compelling argument for such transfer from retained earnings to owners' capital. In addition, McConnell and Muscarella (1985), using a simple random walk model, show that changes in firm value correspond to changes in announced levels of capital expenditure. Furthermore, Trueman (1986) shows that the level of capital expenditure can be a signal about a manager's private information regarding firm value. It is therefore reasonable to believe that the level of capital expenditure is a relevant *ex ante* variable and that investors react to a stock dividend announcement according to their interpretation of this information.

We expect this information to be conditioned by the firm's dividend history. For example, we expect negative market reactions for firms with a bad history and anticipate a greater negative reaction with greater capital expenditures because higher levels of capital expenditure will further exacerbate a cash crunch. Similarly, we expect the positive market reactions to increased capital expenditure by firms with a good history to be larger when the increase is large relative to such firms because a higher level of capital expenditure is expected to result in even larger future cash flows.

More formally,

H3: The announcement period abnormal return is proportional to the level of capital expenditure, increasing for firms with a good history and decreasing for those with a bad history.

Funds from Operations

Some research on stock dividends and splits (Bar-Yosef and Brown 1977) has examined earnings in the post-announcement period as a test of whether announcing firms have outperformed similar non-announcing firms. In contrast, we examine market participants' use of prior information on funds flow in reacting to a stock dividend declaration. The hypothesis is that, conditional on the dividend history, there is an incremental value associated with previously announced information on earnings and funds flow that determines the investor's reaction to a stock dividend announcement.[11] Although the earnings (and funds flow) information is already impounded in the prices prior to a stock dividend announcement, it is part of the *ex ante* information set that investors use to infer the information content of the stock dividend. A history of stable or high levels of funds from operations for a firm announcing a stock dividend corroborates or contradicts the dividend's information content. However, information from just one period prior may not be sufficient since there may be random shocks in the most immediate period. For this reason, we hypothesize that the average funds from operations over the two years immediately preceding the announcement is relevant to investors in revising their expectations about firm value at the time of the stock dividend announcement.[12]

More formally,

H4: The announcement period abnormal return is increasing in the average funds from operations for both types of firms.

II. Data and Sample Selection

Our analysis focuses on firms that declared stock dividends during the calendar years 1976–83. Information on the size of the distribution, previous declarations and the announcement dates are obtained from the CRSP daily master tape. Data on the returns observations during and before the event period for the individual firms in the sample and the market portfolio are obtained from the CRSP daily returns file. Data on firm-specific accounting variables and quarterly earnings announcement dates are obtained from the 1984 annual and quarterly industrial COMPUSTAT files.[13]

We define the event date of the stock distribution as the earlier of (1) the declaration date on the CRSP daily master tape or (2) the announcement date in *The Wall Street Journal*. Day 0 is assumed to be the date on which the market becomes aware of the firm's intention to expand its number of shares.[14] We restricted our sample to stock dividends that were less than or equal to 25 percent to eliminate the effect of a few announcements that were larger than the AICPA (1961) guideline.

We searched *The Wall Street Journal Index* to identify announcement dates and to determine whether a distribution occurred in conjunction with the disclosure of other events. All observations associated with contemporaneous announcements in the seven days before and after the event date were considered contaminated and hence deleted to arrive at a "pure sample." The publication dates of the pure sample were then compared with the cash dividend and stock split dates from the CRSP daily master

---

[11] In a similar approach, Asquith et al. (1989) examine stock splits and the earnings performance of these firms in the pre-announcement period.

[12] The choice of whether to include two, three, or four years is an empirical issue. Going too far back may result in missing the information content, and using the most recent year may result in the use of information on transitory changes.

[13] The COMPUSTAT research tapes that include nonsurviving firms are not used.

[14] For our sample of firms, we found that *The Wall Street Journal* was typically a day later than the CRSP. See Abdel-khalik (1984) for a discussion of the validity of using *The Wall Street Journal* as a source of event dates.

Table 2
Summary of the Sample Selection Process

	1976	*1977*	*1978*	*1979*	*1980*	*1981*	*1982*	*1983*	*Total*
Total stock dividend announcements in each year	223	186	184	176	173	152	126	119	1,339
Less announcements greater than 25%	20	5	2	1	4	4	1	0	37
Less simultaneous cash or stock split announcements	123	101	126	119	108	86	62	49	774
Less announcements by firms not in annual or quarterly COMPUSTAT files	28	21	10	13	20	10	9	7	118
Less simultaneous quarterly earnings announcements	15	17	5	6	5	6	20	14	88
Less other firm-specific announcements in *The Wall Street Journal*	10	16	11	16	13	15	6	4	91
Less multiple announcements by same firm in same financial year	2	4	8	0	2	8	6	8	38
Less announcements by firms with inadequate returns data	0	1	1	0	1	0	0	3	6
Remaining observations	25	21	21	21	20	23	22	34	187

file and the quarterly earnings announcement dates on the quarterly COMPUSTAT file to eliminate the effects of other firm-specific announcements (i.e., confounding events) on the market's response to the stock dividend announcement.

Grinblatt et al. (1984, 465) argued that "by limiting the sample to distributions of 10 percent or more, periodic predictable stock dividends were largely eliminated." We follow a slightly different approach. Instead of choosing a cutoff size, we eliminate firms from our sample that had two or more distributions during the same fiscal year. Also, firms lacking observations for either the 15-day window period (i.e., the prediction period) or the prior period (i.e., the market model estimation period) were deleted from the sample.

Table 2 provides a summary of the sample selection process. Our pure sample consisted of 187 stock dividend announcements between 1976 and 1983 by firms making no other announcements in the 15-day period surrounding the announcement date. As can be seen from table 2, a large number of firms announce their stock dividends simultaneously with other announcements, such as cash dividends.[15] The number of sample

---

[15] There is no significant difference in size of the distribution or industry composition for firms in the sample compared with those in the population. However, the sample firms appear to be somewhat smaller in size relative to the population, presumably because larger firms tend to have more contemporaneous

Table 3
Descriptive Statistics for Stock Dividend Size of the Sample by Year

Year of Announcement	*Number of Announcements*	*Mean Dividend Size*	*Variance of Dividend Size*	*Minimum Dividend Size*	*Maximum Dividend Size*
1976	25	0.0739	0.0043	0.02	0.25
1977	21	0.0714	0.0017	0.03	0.20
1978	21	0.0777	0.0024	0.03	0.20
1979	21	0.0805	0.0018	0.03	0.20
1980	20	0.0835	0.0032	0.03	0.20
1981	23	0.0683	0.0009	0.03	0.10
1982	22	0.0677	0.0010	0.02	0.10
1983	34	0.0741	0.0018	0.02	0.25
All years	187	0.0744	0.0021	0.02	0.25

Note: The stock dividend size corresponds to the FACPR variable in the CRSP daily stock master file, the number of additional shares (per old share) issued as a result of a split or dividend. For example, 0.25 corresponds to a 25% stock dividend. We refer to this variable as DSIZE in our discussion of empirical results.

observations in each year and the average size of the stock dividend for the sample are shown in table 3.

We use the residuals from the one-factor market model (using a value-weighted market index). Parameters of this model are estimated for each sample observation by using ordinary least squares (OLS). In estimating the market model parameters, we use 60 observations from the period immediately preceding the announcement and ending ten days before the announcement date. Further, to consider the possibility that the market anticipated declarations and to investigate its behavior subsequent to declarations, we compute cumulative abnormal returns (CARs) surrounding and including the declaration date.

Summary descriptive statistics for the firms in our sample are given in table 4. The firms with a good history appear to be larger on average and exhibit a smaller relative reduction in retained earnings and a smaller dividend size.

III. Empirical Methodology and Results

Differential Market Response

Table 5 shows the mean seven-day *CAR*s and the associated *t*-statistic for each cell with observations table 1. It can be seen that there are significant differences between cells as well as between dividend history categories. The mean *CAR* is -0.0065 for the single observation of a decrease in both the prior cash dividend and the announced stock dividend. The highest mean *CAR* of 0.0367 (with a *t*-statistic of 4.19) is associated with a maintained stock dividend and no prior cash dividend. The lowest mean *CAR* of

---

announcements. For a more detailed comparison of the sample firms with those in the population in the years under consideration, see Das (1988, 43–50). Das (1992) examines the differential announcement effects and firm characteristics associated with stock splits and dividends that are announced simultaneously with earnings and cash dividends as compared to "pure" announcements. The study demonstrates the corroboration and contradiction effects of contemporaneously disclosed multiple signals.

Table 4
Descriptive Statistics for the Sample Firms

Variable	*Mean*	*Median*	*Standard Deviation*	*Maximum*	*Minimum*
All Firms					
CAR	0.0233	0.0199	0.0747	0.2630	−0.1454
DSIZE	0.0746	0.5000	0.0462	0.2500	0.0200
REC	−0.8486	0.0000	7.1400	0.0000	−95.0896
AVGF	10.9461	4.2975	26.0837	220.3015	−2.3700
CAPEXP	10.4756	2.6560	31.3770	287.0660	0.0000
MKTVAL	57.8739	22.0248	155.8379	1639.0458	2.3331
Bad History Firms					
CAR	−0.0015	−0.0121	0.0913	0.2630	−0.1242
DSIZE	0.0813	0.0550	0.0557	0.2500	0.0200
REC	−2.9770	0.0000	15.4131	0.0000	−95.0896
AVGF	5.3027	4.5120	4.8230	21.0600	−2.3700
CAPEXP	6.0212	2.6665	10.1795	48.1290	0.0330
MKTVAL	28.7781	19.2727	27.7723	154.8162	2.3331
Good History Firms					
CAR	0.0299	0.0259	0.0684	0.2179	−0.1454
DSIZE	0.0728	0.0500	0.0434	0.2500	0.0200
REC	−0.2749	0.0000	0.6869	0.0000	−4.9995
AVGF	12.4669	4.1575	29.1189	220.3015	−1.3685
CAPEXP	11.6762	2.6560	34.8930	287.0660	0.0000
MKTVAL	65.7154	22.3168	174.3034	1639.0458	2.4976

Note: The number of observations used in the construction of this table is 179 because eight firms did not have data on CAPEXP or AVGF. The variables are defined as follows:

CAR = cumulative abnormal returns measured in the 3-day interval before and after the announcement day,
DSIZE = the size of the distribution measured as a fraction. It corresponds to the *FACPR* variable in CRSP,
REC = the value of the retained earnings constraint and is measured as described in the appendix,
CAPEXP = the value of capital expenditure (annual COMPUSTAT data item #128) in the fiscal year immediately preceding the announcement year (measured in millions of dollars),
AVGF = the average value of funds from operations (annual COMPUSTAT data item #110) in the two fiscal years immediately preceding the announcement year (measured in millions of dollars), and
MKTVAL = the closing price times the number of shares outstanding at the beginning of the fiscal year in which an announcement is made (measured in millions of dollars).

−0.0324 (t-statistic = 1.75) is associated with the 11 firms that discontinued the cash dividend and had no prior stock dividend.

In contrast to this negative market reaction following substitution of stock for cash dividends, a positive (although statistically indistinguishable from zero) reaction followed the nine announcements of a discontinued cash dividend with a maintained level of stock dividend. To explore this seemingly counterintuitive result, we examined some of the prior accounting information associated with these observations and found that the mean and median levels of capital expenditure and funds from operations (after

Table 5
Mean Market Reactions by Dividend History Category

Cash Dividend History	*Announced Stock Dividend*			
	Decreased	*Maintained*	*No Prior Dividend*	*Total*
Discontinued				
Observations	1	9	11	21
CARs	−0.0065	0.0134	−0.0324	−0.0115
(*t*-statistic)	(−)	(0.345)	(−1.745)	(−0.596)
Maintained				
Observations	5	41	9	55
CARs	0.0258	0.0295	−0.0045	0.0236
(*t*-statistic)	(0.574)	(3.537)	(−0.216)	(2.917)
No Prior				
Observations	12	72	27	111
CARs	0.0049	0.0367	0.0288	0.0313
(*t*-statistic)	(0.173)	(4.186)	(2.153)	(4.352)
Total				
Observations	18	122	47	187
CARs	0.0100	0.0326	0.0081	0.0243
(*t*-statistic)	(0.459)	(5.022)	(0.791)	(4.484)

Note: The 7-day cumulative abnormal returns (CARs) represent the mean market reaction beginning with day −3 and ending on day +3, inclusive of the announcement day.

adjusting for size) are higher for the nine firms that maintained stock dividends relative to the 11 firms that initiated stock dividends (see table 6). This is consistent with the argument that the former group of firms discontinued cash dividends to finance more profitable investments while the latter group discontinued cash dividends because of lack of funds. This argument is further supported by the fact that the 11 firms that initiated stock dividends had, on average, a more binding constraint on retained earnings, whereas the nine firms that maintained a prior stock dividend after discontinuing cash dividends had relatively large levels of unrestricted retained earnings.

We also test the null hypotheis of no association between the sign of the abnormal return in the announcement period and the firm's dividend history. Using a 2×2 classification table (see table 7), we compare the observed sign of the market reaction with the dividend history category of the announcing firm. The chi-square test (which assumes independence) rejects the null hypothesis and indicates a heterogeneous population. In other words, there is evidence of association between the sign of the market reaction and the dividend history category of the announcing firm.

Another way to examine differences in returns between firms with a good and bad history is to examine the market reaction of each category over a period before and after the announcement. We thus accumulate the mean abnormal returns within three portfolios of firms aggregated as (1) the total sample, (2) firms with good history, and (3) firms with bad history. The plots of the magnitude of the CAR against time (day) are

Table 6
Comparison of Selected Firm-Specific Variables for Firms Discontinuing Cash Dividends

	Funds From Operations (per $ of firm value)	*Capital Expenditure (per $ of firm value)*	*Retained Earnings Constraint*
Stock dividend initiated (*N*=11)			
Mean	0.171	0.132	−9.08
Median	0.139	0.079	−0.09
Stock dividend maintained (*N*=9)			
Mean	0.277	0.277	−0.53
Median	0.269	0.178	0

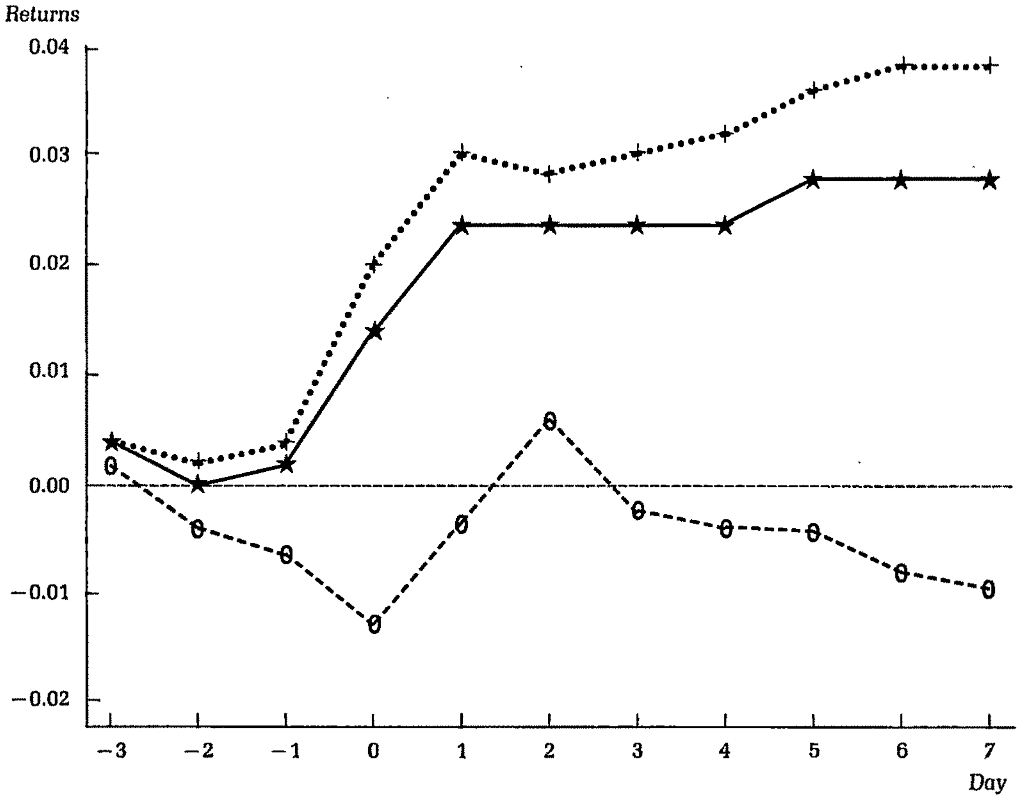
Table 7
Chi-Square Table for Dividend History Category and Market Reaction

Dividend History Category	*Observed CAR*		
	Negative	*Positive*	*Total*
Bad History			
Frequency	23	15	38
Expected Frequency	14.20	23.80	
Good History			
Frequency	47	102	149
Expected Frequency	55.80	93.20	
All Firms			
Frequency	70	117	187

Note: The chi-square statistic=10.859 (p-value=0.001, d.f., 1); the likelihood ratio chi-square statistic=10.547 (p-value=0.001, d.f., 1).

shown in figure 1, where day 0 is the announcement day. The pattern for the sample as a whole reflects that for the firms with a good history since the sample is dominated by that group by a 4:1 ratio. However, it is apparent from figure 1 that there are indeed two distinct categories within the population of firms declaring stock dividends. The mean seven-day *CAR* for the firms with a good history was 0.0308 with a *t*-statistic of 5.56 (which rejects the null of being equal to 0 at a p-value of 0.0001). The mean *CAR* for the firms with a bad history was −0.0015 with a *t*-statistic of −0.10 and a p-value of 0.9227 (which does not allow rejection of the null). Average market reaction for the firms of good history was significantly different from that for those with a bad history, as evi-

Figure 1
CAR Plot



Note: Day 0 is the stock dividend announcement day.

••• + ••• Good history (n = 149)
---0--- Bad history (n = 38)
—★— All firms (n = 187)

denced by the Mann-Whitney test (14,634 with a p-value of 0.007) and the Welch's t-statistic (1.97 with a p-value of 0.05). The mean market reaction of 0.0243 across all firms was statistically significant, with a t-statistic of 4.48 (table 5).

Multiple Regression: Explaining Cross-Sectional Variations

Our hypotheses suggest that the market uses information other than the dividend history to corroborate or contradict the information conveyed by a stock dividend. In this section, we report on a multiple regression approach to examine cross-sectional variations in market response to test our hypotheses. A correlation matrix of the variables is provided in table 8 to facilitate interpretation of the results. As seen in table 8, capital expenditures (CAPEXP) and funds flow from operations (AVGF) are significantly correlated with dividend size (DSIZE) and the retained earnings constraint

Table 8
Correlation Matrix of Variables

	DSIZE	*REC*	*AVGF*	*CAPEXP*
DSIZE	—	−0.0273 (0.72)	−0.1301 (0.08)	−0.1333 (0.08)
REC	−0.3075 (0.00)	—	0.0391 (0.60)	0.1333 (0.07)
AVGF	−0.0294 (0.69)	0.0361 (0.63)	—	0.7786 (0.00)
CAPEXP	0.0130 (0.86)	0.0078 (0.92)	0.7794 (0.00)	—

Note: Numbers above the diagonal represent Spearman rank correlations while those below the diagonal represent Pearson correlations. The numbers in parentheses represent *p*-values. Variable names and measures are defined in table 4.

(REC). This is not surprising because REC is derived by using the DSIZE variable.

We first estimate the following regression equation:

$$CAR_i=a+\sum_{k=1}^{4}b_kX_{ik}+e_i; \qquad i=1,\ldots,N. \qquad (1)$$

Here CAR_i is the seven-day cumulative abnormal return in the announcement period corresponding to the *i*th firm. The variables X_{ik} are the four firm-specific accounting variables: DSIZE, REC, CAPEXP, and AVGF. This specification resembles that of Grinblatt et al. (1984). In conformance with prior research, all firms are treated as homogeneous with respect to prior information except for cross-sectional variation in the included independent variables.

Previous research explaining cross-sectional variations in abnormal returns with firm-specific variables has argued that the variance of the residuals may be greater for small firms (see Atiase 1985; Banz 1981; Zeghal 1984). We examined the residuals from the OLS estimation of equation (1) for heteroscedasticity by using the direct test of Park and Glejser (Pindyck and Rubinfeld 1981, 150). On regressing the absolute value of the residuals from the OLS regression against the square root of the inverse of the size variable (beginning market value), we found strong evidence of heteroscedasticity.[16] We therefore reestimated equation (1) with weighted least squares (WLS) using the beginning-of-period market value of a firm as the weight.[17] The results of the WLS regression are given in table 9, and comparative results from OLS estimation are given in table 10. The F-statistic is higher with WLS estimation (14.78, $p=0.0001$) than with OLS estima-

---

[16] The R^2 values of the regression was 54 percent, and the *t*-statistic for the size variable was 14.44.

[17] Beginning market value is measured as the closing price times the number of shares outstanding at the beginning of the fiscal year in which an announcement is made; see Christie (1987).

Table 9
Results of WLS Multiple Regression
(with and without a dummy variable for prior dividend history)

	Independent Variables (t-statistic)												
	Intercept	d_t	*DSIZE*	*N-DSIZE*	*REC*	*N-REC*	*CAPEXP*	*N-CAPEXP*	*AVGF*	*N-AVGF*	F-*statistic* (*p-value*)	R^2	*Adjusted* R^2
Model 1 (eq. [1])	−0.0038	—	0.3474	—	0.0016	—	0.0004	—	−0.0003	—	14.78	0.25	0.24
	(−0.47)	—	(3.62)	—	(2.61)	—	(4.47)	—	(−3.11)	—	(0.0001)		
Model 2 (eq. [2])	−0.0061	—	0.4638	−0.6047	0.0024	−0.0025	0.0004	−0.0043	−0.0003	0.0063	10.73	0.34	0.30
	(−0.76)	—	(4.72)	(−3.14)	(0.25)	(−0.27)	(4.31)	(−3.05)	(−3.37)	(2.51)	(0.0001)		
Model 2A (DSIZE omitted)	0.0189	—	—	—	−0.0034	0.0039	0.0005	−0.0039	−0.0004	0.0026	9.01	0.24	0.21
	(3.11)	—	—	—	(−0.34)	(0.39)	(6.38)	(−2.59)	(−4.37)	(1.09)	(0.0001)		
Model 2B (REC omitted)	−0.0065	—	0.4614	−0.5769	—	—	0.0004	−0.0043	−0.0003	0.0061	14.45	0.34	0.31
	(−0.81)	—	(4.75)	(−3.74)	—	—	(4.43)	(−3.10)	(−3.40)	(2.67)	(0.0001)		
Model M2C (CAPEXP omitted)	0.0139	—	0.5048	−0.3875	0.0047	−0.0040	0.0001	−0.0014	—	—	10.06	0.26	0.23
	(−1.90)	—	(5.13)	(−2.13)	(0.48)	(−0.41)	(2.47)	(−1.36)	—	—	(0.0001)		
Model 2D (AVGF omitted)	−0.0141	—	0.6178	−0.6907	0.0092	−0.0088	—	—	0.0000	0.0014	8.53	0.23	0.20
	(−1.71)	—	(6.43)	(−3.40)	(0.92)	(−0.88)	—	—	(0.24)	(0.80)	(0.0001)		
Model 3 (Intercept plus slope dummies)	−0.0061	−0.0003	0.4637	−0.6029	0.0024	−0.0025	0.0004	−0.0044	−0.0003	0.0063	9.48	0.34	0.30
	(−0.73)	(−0.01)	(4.62)	(−2.31)	(0.25)	(−0.27)	(4.29)	(−2.83)	(−3.34)	(1.76)	(0.0001)		

Note: The number of observations in each regression is 179. The variable definitions are given in table 4; variables with an N-prefix represent the original variables multiplied by the dummy variable, d_t. Beginning market value is the weight for the WLS regressions.

Table 10
Results of OLS Multiple Regression
(with and without a dummy variable, d_t, for prior dividend history)

	Independent Variables (t-statistic)												
	Intercept	d_t	DSIZE	N-DSIZE	REC	N-REC	CAPEXP	N-CAPEXP	AVGF	N-AVGF	F-*statistic* (p-*value*)	R^2	*Adjusted* R^2
Model 1 (eq. [1])	0.0085 (0.77)	— —	+0.2061 (1.62)	— —	0.0014 (1.69)	— —	0.0003 (1.09)	— —	−0.0002 (−0.73)	— —	1.39 (0.24)	0.03	0.01
Model 2 (eq. [2])	0.0085 (0.75)	— —	0.3179 (2.36)	−0.4126 (−2.16)	0.0061 (0.68)	−0.0059 (−0.65)	0.0004 (1.43)	−0.0035 (−1.98)	−0.0004 (−1.13)	0.0036 (1.06)	2.22 (0.03)	0.09	0.05
Model 2A (DSIZE omitted)	0.0093 (0.85)	— —	0.2956 (2.24)	−0.3136 (−1.86)	0.0065 (0.72)	−0.0060 (−0.66)	0.0001 (0.84)	−0.0019 (−1.71)	— —	— —	2.58 (0.02)	0.08	0.05
Model 2B (REC omitted)	0.0099 (0.87)	— —	0.3099 (2.29)	−0.3662 (−1.92)	0.0071 (0.78)	−0.0065 (−0.71)	— —	— —	−0.0000 (−0.06)	−0.0013 (−0.62)	2.06 (0.06)	0.07	0.03
Model 3 (Intercept plus slope dummies)	0.0083 (0.67)	0.0011 (0.04)	0.3195 (2.23)	−0.4206 (−1.43)	0.0061 (0.67)	−0.0059 (−0.65)	0.0004 (1.42)	−0.0035 (−1.91)	−0.0004 (−1.12)	0.0035 (0.90)	1.96 (0.05)	0.09	0.05

Note: The number of observations in each regression is 179. The variable definitions are given in table 4; variables with an N-prefix represent the original variables multiplied by the dummy variable, d_t.

tion (1.39, $p=0.24$), and all of the coefficients except the intercept are statistically significant.[18]

The above specification assumes, as does prior research, that the population of firms announcing stock dividends is homogeneous. However, if previously disclosed accounting information affects the market response to the stock dividend announcement, then the coefficient estimates associated with the corresponding independent variables would be different across the two types of firms. We therefore specify the following regression equation:

$$CAR_i=a'+\sum_{k=1}^{4}b_k'X_{ik}+\sum_{k=1}^{4}c_k(X_{ik}*d_i)+e_i', \quad i=1,2,\ldots,N, \tag{2}$$

where d_i is a dummy variable that takes on a value of 1 for firms with a bad history and 0 otherwise. WLS estimates for the above regression are given as model 2 in table 9, where the prefix N indicates variables associated with the dummy variables.

To examine the heterogeneity of the sample firms, we first test the null hypothesis that the coefficients associated with the dummy variables are all zero for firms with a bad history. The F-test rejects the null at $p=0.0005$, which indicates that the model has greater explanatory power when dummy explanatory variables are used to condition the accounting information effects on the dividend history.

The coefficient estimates in table 9 (model 2) are statistically different from zero for the dividend size variable. We observe a differential market reaction to firms on the basis of their dividend history. For firms with a good (bad) history, abnormal returns increase (decrease) with the size of the dividend. For firms with a bad history, the value of the coefficient is -0.1409 [$=(0.4638)+(-0.6047)$], and an F-test to evaluate whether the coefficients are different for these firms is significant at the 0.002 level ($F=9.89$). Note, however, that not conditioning on dividend history (model 1) would lead to the conclusion that the market reaction is always increasing in the size of the stock dividend distribution for all firms.

As table 9 (model 2) indicates, the coefficient corresponding to the retained earnings constraint is negative for firms with a bad history and positive for firms with a good history. The signs are consistent with our hypothesis H2, but the coefficient estimates in both cases are not significantly different from zero.[19] We need to interpret the coefficient estimates of REC and DSIZE together since these variables are correlated. We reestimate our regression equation (2), eliminating either the REC or DSIZE variable from the model. These results are reported as models 2A and 2B in table 9. When the DSIZE variables are omitted (model 2A), the REC variable for the good history firms is insignificant and negative, whereas it is insignificant and positive for bad history firms. When the REC variables are omitted (model 2B), in conformance with our hypothesis, the coefficients for DSIZE are significantly positive for firms with a good history and significantly negative for those with a bad history. The magnitude of the coefficient in model 2B changes marginally relative to the estimates obtained from the full model 2.

---

[18] In an earlier analysis, the independent variables for CAPEXP and AVGF were computed per dollar of firm value. Because an examination of the error structure showed that the errors were heteroscedastic, both the dependent and independent variables were further deflated by the beginning-of-period market value. The results obtained were very similar to the conclusions reported here.

[19] Setting $REC_i=-RED_i/RE_i$ for all firms, where RE_i is retained earnings (COMPUSTAT item #36) did not materially change the coefficients of other variables. The sign of the coefficients on REC_i and $N\text{-}REC_i$ changed but remained insignificant. Similarly, replacing URE_i in the denominator by RE_i only for firms that did not report unrestricted retained earnings (instead of setting them equal to zero) also did not alter the basic results.

The results for CAPEXP are statistically significant for both groups of firms. For those with a good history, market reaction increases with the level of CAPEXP, but is significantly smaller for those with a bad history. Indeed, for those firms, a negative association indicated by a coefficient of -0.0039 [$=(0.0004)+(-0.0043)$] suggests that as the size of the capital expenditure increases, the worse is the market reaction. Moreover, an F-test shows that the coefficients for the bad history firms are different from those for good history firms and are significant at the 0.003 ($F=9.29$) level. This result is consistent with our hypothesis H3. Note that, without conditioning on dividend history (model 1), the market reaction is increasing in the size of the capital expenditure. Partitioning the set of firms clearly shows that the market uses complementary sources of information in a differential manner.

Results for average funds from operations indicate a negative sign for firms of good history and a positive sign for firms of bad history which appears to be contrary to one part of our hypothesis H4. This contradiction is likely influenced by collinearity among the independent variables in the regression. The funds flow and capital expenditure variables are highly correlated (correlation coefficient=0.78; see table 8). Collinearity diagnostics also indicate a high condition index of 44.3, which is well above the upper bound of 30 set by Belsey et al. (1980). We therefore reestimated our regression equation, eliminating in turn either the CAPEXP or the AVGF variables. These results are reported as models 2C and 2D in table 9. When the CAPEXP variables are omitted, both AVGF and N-AVGF have a positive sign (which is consistent with hypothesis H4), but neither is significantly different from zero. When the average funds flow variables are omitted, coefficients for both CAPEXP and N-CAPEXP change only marginally from the estimates obtained from the full model. In both cases, the coefficients for the remaining variables do not change appreciably.

IV. Concluding Remarks

Our results indicate a systematic relation between the magnitude of the abnormal returns at the time of the stock dividend announcement and previously disclosed accounting information, and this relation is conditioned by the dividend history of the firm. In particular, the same previously disclosed accounting information is perceived differentially by investors according to the good or bad history of the announcing firm. More importantly, this research documents the complementarity of various forms of firm-specific information that is used by investors to interpret subsequent disclosures. We acknowledge, however, that the small number of observations for firms of bad history limits the statistical power of our analysis, as is shown in model 3. Allowing for an intercept dummy in addition to the slope dummies is equivalent to estimating two separate regressions for the two types of firms. The estimated intercept terms are almost identical for both groups, and the other coefficient estimates are also similar to those in model 2.

From an accounting policy perspective, it is important to identify conditioning variables that are useful in predicting differential effects across firms. It is important for policy makers to take these complementarities into account when deciding on disclosure issues.[20] Reliance on contemporaneous market association tests as the sole determinant of the value of disclosure is myopic. Our analysis suggests that accounting disclosures can be value-relevant as complementary sources of information, long after they are published. Research on how such multiple signals are used by investors in reacting to disclosures other than stock dividend announcements will likely further our understanding of the way investors value firms.

Appendix

Measure of Retained Earnings Constraint

The variable of interest is the relative reduction in unrestricted retained earnings. Data item #97 in the annual industrial COMPUSTAT file gives the dollar amount of retained earnings that are not restricted by debt agreements (URE_i). Since a stock dividend entails a decrease in the retained earnings of the firm, the relative reduction can be expressed as a ratio of the amount by which the unrestricted retained earnings will be reduced subsequent to an announcement. More formally, we can represent this as follows:

$$\text{Retained Earnings Constraint: } REC_i = -\frac{RED_i}{URE_i},$$

where, for firm i, RED_i is the reduction in the book value of retained earnings following the announcement of a stock dividend, and represents the fair market value of stock transferred to owners' capital. RED_i is measured as the size of the distribution times the market value of shares outstanding at the beginning of the announcement quarter. URE_i is the level of unrestricted retained earnings prior to the announcement. For firms that did not report unrestricted retained earnings, REC_i is set equal to zero. The assumption is that firms that did not disclose unrestricted retained earnings did not have debt covenants restricting their retained earnings. Similar assumptions were made by Bowen et al. (1981), Lys (1984), and Shevlin (1987). In our sample of 179 firms, 90 did not report the level of unrestricted retained earnings.

---

[20] The Report of the Advisory Committee on Corporate Disclosure (SEC 1977, 271–72) states that "knowledge about the importance to investors of a variety of information types could assist the Advisory Committee in its determinations and recommendations concerning the areas of corporate disclosure which may need emphasis (or de-emphasis)."

References

Abdel-khalik, A. R. 1984. A note on the validity of the WSJ as a source of event dates. *Journal of Accounting Research* 22 (Autumn): 758–59.

Aharony, J., and I. Swary. 1980. Quarterly dividend and earnings announcements and stockholders' returns: An empirical analysis. *Journal of Finance* 35 (March): 1–12.

American Institute of Certified Public Accountants. 1961. *Accounting Research and Terminology Bulletins.* Final Edition. New York: AICPA.

Antle, R., J. S. Demski, and S. G. Ryan. 1991. Accounting earnings with multiple sources of information. Working paper, Yale University, New Haven, CT.

Asquith, P., P. Healy, and K. Palepu. 1989. Earnings and stock splits. *The Accounting Review* 64 (July): 387–403.

Atiase, R. 1985. Predisclosure information, firm capitalization and security price behavior around earnings announcements. *Journal of Accounting Research* 23 (Spring): 21–36.

Banz, R. W. 1981. The relationship between return and market value of common stock. *Journal of Financial Economics* 9 (March): 3–18.

Bar-Yosef, S., and L. Brown. 1977. The earnings information content of stock distribution announcements. In *Proceedings of the American Accounting Association Annual Meeting*. Sarasota, FL: AAA.

Belsey, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data Sources of Collinearity.* New York: John Wiley & Sons.

Bowen, R. M., E. W. Noreen, and J. M. Lacey. 1981. Determinants of the corporate decision to capitalize interest. *Journal of Accounting & Economics* 3 (August): 151–79.

Christie, A. A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting & Economics* 9 (September): 231–58.

Das, S. 1988. Differential information content of stock dividend announcements. Ph.D. dissertation, Carnegie Mellon University, Pittsburgh, PA.

———. 1992. Concurrent discretionary and routine disclosures: Evidence from earnings, dividends and stock distributions. Working paper, University of California at Berkeley.

Davidson, S., C. Stickney, and R. Weil. 1988. *Financial Accounting: An Introduction to Concepts, Methods, and Uses*. 4th ed. Hinsdale, IL: Dryden Press.

Dielman, T. E., and H. R. Oppenheimer. 1984. An examination of investor behavior during periods of large dividend changes. *Journal of Financial and Quantitative Analysis* 19 (June): 197–216.

Eisemann, P. C., and E. A. Moses. 1978. Stock dividends: Management's view. *Financial Analyst Journal* 34 (July-August): 77–83.

Fama, E., L. Fisher, M. Jensen, and R. Roll. 1969. The adjustment of stock prices to new information. *International Economic Review* 10 (February): 1–21.

Gonedes, N. 1978. Corporate signalling, external accounting, and capital market equilibrium: Evidence on dividends, income, and extraordinary items. *Journal of Accounting Research* 16 (Spring): 26–79.

Grinblatt, M. S., R. W. Masulis, and S. Titman. 1984. The valuation effects of stock splits and stock dividends. *Journal of Financial Economics* 13 (December): 461–90.

John, K., and L. H. P. Lang. 1991. Insider trading around dividend announcements: Theory and evidence. *Journal of Finance* 46 (September): 1361–89.

Kalay, A. 1980. Signalling, information content, and the reluctance to cut dividends. *Journal of Financial and Quantitative Analysis* 15 (November): 855–69.

Kane, A., Y. K. Lee, and A. Marcus. 1984. Earnings and dividend announcements: Is there a corroboration effect? *Journal of Finance* 39 (September): 1091–99.

Lakonishok, J., and B. Lev. 1987. Stock splits and stock dividends: Why, who, and when. *Journal of Finance* 42 (September): 913–32.

Leland, H. E., and D. H. Pyle. 1977. Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance* 32 (May): 371–87.

Lys, T. 1984. Mandated accounting changes and debt covenants: The case of oil and gas accounting. *Journal of Accounting & Economics* 6 (April): 39–65.

McConnell, J. J., and C. Muscarella. 1985. Corporate capital expenditure decisions and the market value of the firm. *Journal of Financial Economics* 14 (September): 399–422.

McNichols, M., and A. Dravid. 1990. Stock dividends, stock splits, and signalling. *Journal of Finance* 45 (July): 857–79.

Ou, J. A., and S. H. Penman. 1989. Financial statement analysis and the prediction of stock returns. *Journal of Accounting & Economics* 11 (November): 295–329.

Pindyck, R. S., and D. Rubinfeld. 1981. *Econometric Models and Economic Forecasts*. 2d ed. New York: McGraw-Hill.

Securities and Exchange Commission. 1977. *Report of the SEC Advisory Committee on Corporate Disclosure*. Washington, D.C.: U.S. Government Printing Office.

Shefrin, H. M., and M. Statman. 1984. Explaining investor preference for cash dividends. *Journal of Financial Economics* 13 (June): 253–82.

Shevlin, T. 1987. Taxes and off-balance sheet financing: Research and development limited partnerships. *The Accounting Review* 62 (July): 425–42.

Sosnick, S. H. 1961. Stock dividends are lemons, not melons. *California Management Review* 3 (Winter): 61–82.

Talmor, E. 1981. Asymmetric information, signaling, and optimal corporate financial decisions. *Journal of Financial and Quantitative Analysis* 16 (November): 413–38.

Trueman, B. 1986. The relationship between the level of capital expenditures and firm value. *Journal of Financial and Quantitative Analysis* 21 (June): 115–29.

Woolridge, R. J. 1983. Stock dividends as signals. *The Journal of Financial Research* 6 (Spring): 1–12.

Zeghal, D. 1984. Firm size and informational content of financial statements. *Journal of Financial and Quantitative Analysis* 19 (September): 299–310.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 48–65

Agency and Efficiency in Nonprofit Organizations: The Case of "Specific Health Focus" Charities

Jeffrey L. Callen
The Hebrew University and
New York University
Haim Falk
McMaster University

SYNOPSIS AND INTRODUCTION: This article relates the efficiency of nonprofit organizations to the composition of their board of trustees. Following arguments by Fama and Jensen (1983a, 1983b) and Williamson (1983), we conjecture that nonprofit organizations are more efficient if their board of trustees have a larger proportion of outsider trustees rather than insider (employee) trustees since the former are presumed to have a greater incentive to monitor the organization and the latter are presumed to have a greater incentive to consume perquisites. This conjecture is tested empirically for a sample of 72 charities with a specific health focus, with respect to both the technical and allocative efficiency of the organization.[1]

Technical efficiency indices were derived by using data envelopment analysis (DEA) under various assumptions concerning the industry production technology, including the assumption that charity outputs are not substitutable for each other. These indices are reasonably well measured since the data included information on volunteer labor, a necessary input for charity services.

To test the conjecture that the technical efficiency of the organization is affected by the composition of the board of trustees, the estimated

---

[1] The distinction between technical and allocative efficiency is from Farrell (1957). The literature on efficiency measurement is quite voluminous (see the references in section III). Several attempts to rank nonprofits, including charities, by relative efficiency are documented in the literature (Clotfelter and Salamon 1982; Hochman and Rodgers 1977; Rose-Ackerman 1982; Steinber 1986; Strnat 1986; Weisbrod 1980; Weisbrod and Schlesinger 1986), but these studies did not distinguish between technical and allocative inefficiencies.

We acknowledge the input of anonymous referees, and of the seminar participants at the Hebrew University, McMaster University, and New York University, especially Ajay Maindiratta, for comments on an earlier draft. The assistance of Dading Li and Rita Ruffo is also appreciated. Jeffrey Callen acknowledges financial support from the Accounting Research and Education Centre, McMaster University, and the Krueger Center, The Hebrew University.

Submitted August 1991.
Accepted August 1992.

indices of technical efficiency were regressed on the board's proportion of insider trustees and the organization's debt-value ratio. The conjectured relationship was not confirmed by the data. As an additional test, the proportion of insider trustees was replaced by the ratio of insider-trustee remuneration to total labor remuneration, the argument being that the greater the relative remuneration received by the insider trustees, the more capable they might be in appropriating perquisites of all types. As before, the results did not confirm the conjectured relationship, leading us to conclude that the technical efficiency of nonprofits is not affected by the composition of the board of trustees.

To determine the relationship, if any, between the allocative efficiency of the organization and the composition of the board of trustees, charity costs were multiplied by the technical efficiency index so that any remaining cost inefficiencies are allocative in nature. This adjusted cost was then regressed on the proportion of insider trustees on the board (or the proportion of insider-trustee remuneration) and the organization's debt-value ratio. The conjectured relationship between allocative efficiency and the composition of the board of trustees was also not confirmed by the data.

Key Words: *Nonprofit organizations, Efficiency, Board of trustees, Data envelopment analysis.*

Data Availability: *The data used in this study are available from the authors.*

THE remainder of this article is organized as follows. Section I presents the theoretical background and motivation for this study. Section II develops the empirical research design; section III briefly describes technical efficiency indices; section IV describes the data. The empirical results are reported in section V, and section VI concludes the article. Technical efficiency indices for nonsubstitutable output technologies are derived in the appendix.

I. Background and Motivation

Although privately organized nonprofit charities play an important role in the economy, their economic structure is not well understood. Several attempts to explain the existence of privately organized charities are offered in the literature.[2] In one such explanation, Weisbrod (1975) proposes that private charity organizations efficiently convert gifts into services and serve as a conduit between their donors and their clients. This raises the question of why private charities should be more efficient in providing their services than alternative forms of organization.

---

[2] For example, it has been argued that private charities arose in response to contract failure (Easley and O'Hara 1986; Hansman 1980; Pauly 1980; Weisbrod 1980) and government failure (Niskanen 1971; Thompson 1980; Weisbrod 1980). Others (Feldstein and Clotfelter 1976; Whinston 1980) view charities in the context of social contracts and norms. For an extensive review of the literature and some implications for accounting, see Falk (1992).

Fama and Jensen (1983a, 1983b) argue that an initial assessment would be to assume that nonprofit organizations are generally inefficient by comparision to profit-seeking organizations, which distribute profit surpluses to shareholders. Shareholders have an incentive to monitor management activities and to motivate managers to reduce inefficiencies, thereby increasing surpluses. Nonprofit organizations, in contradistinction, do not have shareholders *per se* since rights over the surplus generated by a nonprofit organization are not alienable. Hence, the proper incentives to monitor managerial activities and reduce wastage in the nonprofit organization are likely to be absent.

Notwithstanding this argument, Fama and Jensen maintain that, in fact, the nonprofit organizational form is optimal for charitable activities because charities, unlike profit-seeking institutions, have access to low cost capital, namely, donor capital. Donors are willing to supply this form of capital to charities for social, psychological, religious, and agency cost reasons, but are unwilling to donate capital to profit-seeking institutions except in unusual circumstances. Furthermore, Fama and Jensen argue that not only does donor capital replace equity capital in charities, but donors, especially large donors, also displace shareholders in monitoring management activities. Although there are no surpluses to share in a charity, large donors likely require a (nonpecuniary) return from providing capital. This return may be altruistic in nature; for example, the satisfaction that the charity's activities are being provided efficiently or, less altruistically, the return may be public acclaim and the promotion of business activities and personal reputation. Be that as it may, there is probably little return from association with an inefficient, wasteful charity, especially if such wastefulness becomes known to the public at large.[3] This concern motivates large donors to monitor the charity's efficient use of resources.

Williamson (1983) takes this argument one step further by observing that monitoring management is costly. Moreover, the monitoring function in charities is bound to be more complicated than for business enterprises, and hence more costly, because "ownership" is not negotiable and market prices are unavailable to assess the periodic "added value" produced by management. Also, some charity inputs and outputs are qualitative in nature or gauged by measures other than money (e.g., hours of volunteer work, etc.). Thus, traditional periodic financial reports that are limited to monetary information are not a sufficient monitoring device (even when independently audited). Donors' participation on the charity's major decision making body, namely the board of trustees, may be a relatively cheap way to monitor management activities effectively.

Although these arguments are fairly compelling, they are conjectural nevertheless. One could easily argue that donors do not have the motivation to monitor management activities to the same extent that shareholders do. Since the return to donor capital is nonmonetary, donors do not have a direct and immediate incentive to attend to dysfunctional managerial behavior. Also, as Williamson (1983) has pointed out, large donors may take their responsibilities in a fairly casual manner or may have their own hidden agendas. For example, donors may be more concerned with praise and publicity than with the efficient allocation of resources or the ultimate disposition of their donations.

---

[3] The recent scandal concerning United Way of America, the largest U.S. charity organization, is a case in point. It was alleged that the 20-year president received an overall annual compensation package of $464,000, enjoyed opulent perquisites, and was engaged in lavish spending and patronage (Shepard 1992). Despite exoneration by a preliminary independent review by an agency hired by the board, the president was forced to resign by public opprobrium and pressure from the United Way chapters (Babcock 1992; Miller et al. 1992).

In addition, it may be that insider employees (management) and volunteers are the ones who truly monitor the efficient allocation of resources in charities. In fact, the labor market for charities may be characterized by self-selection in the sense that those who work for charities do so because they empathize with and are committed to the goals of the organization. As a consequence, it can be argued that employees, *a fortiori* volunteers, are more motivated to perform the monitoring function than donors. It is also plausible that debtholders, and other claimants on the organization, help to monitor charity activities. Governmental oversight of charities (audits) may also curb inefficiencies, as noted by Williamson.

The conjecture that large donors effectively monitor management activities in charities is tested empirically in our study of Canadian charities with a specific health focus. It should be noted at the outset, however, that our tests are not consistent with the arguments expressed by Fama and Jensen (1983a, 1983b) or Williamson (1983). Fama and Jensen emphasize the relationship between organizational efficiency and major money donors. Williamson, who generalizes from large donors to outsider trustees, also emphasizes the monetary contribution of outsiders. Although we were able to obtain fairly complete data on the composition of trustee boards, we could not obtain useful data concerning individual contributions. Instead, we employ proxy variables for major donors, such as the ratio of outsider to total trustees and (one minus) the debt-value ratio, but recognize that there is no one-to-one relationship between these variables and the number of major donors. In fact, outsider trustees may be those willing to donate time and expertise rather than money.

Even if our approach is somewhat indirect, our empirical results are useful in answering several questions concerning the governance of nonprofit organizations. What is the relationship between organizational efficiency and the composition of the board of trustees? Does the composition of the board of trustees affect primarily technical efficiency or allocative efficiency? Is there any relationship between organizational efficiency and the debt-value ratio of nonprofits? To shed light on these issues, we investigate the relationship between organizational efficiency and the composition of boards of directors in nonprofit organizations.

II. Research Design

According to Williamson (1983, 359) the major testable hypothesis to be derived from the Fama-Jensen framework concerns the relationship between expenses per dollar's worth of receipts and the proportion of insider employees (compared to outsider donors) who are executive officers, directors, or trustees (hereinafter trustees) of the charity.

> Fama and Jensen . . . have the beginnings of a predictive theory of the performance of nonprofits based on the composition and character of the board of directors (trustees). For example, I would predict on their reasoning that the expense-to-payout ratio will be higher in the nonprofits where the proportion of insiders on the board is high and where outsiders make negligible contributions.

The reasoning is that outsider trustees are generally large donors and that large donors are properly motivated to monitor management activities in the organization. By contrast, insider trustees, being managers of the organization, have the opposite incentive, namely to allocate themselves larger salaries, larger offices, and other perquisites. Therefore, if the Fama and Jensen framework is correct, there should be a direct relationship between charity expenses and the proportion of insider to outsider trustees. However, if large donors are unconcerned about efficiency of charitable organizations, or employee trustees as well as donors are equally concerned about charity efficiency, no relationship will be observed between charity expenses and the proportion of insider to outsider trustees, which would refute the Fama-Jensen-Williamson (FJW) conjecture.

Charity costs are a function of the particular factor input-output combinations chosen by the institution, as well as donor-monitoring effectiveness (if any). Different charities are bound to operate with different input-output mixes, if only because of differential access to volunteer labor. These differences have to be accounted for prior to testing the FJW conjecture.[4]

Even if the effect of the input-output mix on costs is accounted for, estimating the relationship between charity expenses and the proportion of insider to total trustees is not sufficient to determine the nature of the cost inefficiency. The economics literature has differentiated between technical and allocative inefficiencies. An organization is technically inefficient if it is simply wasting resources; that is, it could be producing more with the given level of resources, or produce the same amount with fewer resources. In other words, the organization is technically inefficient if it is not producing on the production frontier irrespective of relative prices. Yet, even if the organization is operating on its production frontier, it may not be using the appropriate input mix, given relative input prices. This latter phenomenon is called allocative inefficiency.

This dichotomy between technical and allocative efficiency leads to a conceptually similar dichotomy of the FJW conjecture: (1) donor-monitoring activities mitigate against technical inefficiencies in charities and (2) donor-monitoring activities mitigate against allocative inefficiencies. It should be noted, however, that these two facets are neither mutually exclusive nor inclusive. Donors may be attenuating both technical and allocative inefficiencies. Alternatively, donors may be monitoring one type of inefficiency and not the other. Indeed, intuition suggests that donor trustees are more likely concerned with broad policy decisions concerning the charity's direction than with day-to-day operations and management's excess consumption of perquisites. Such strategic decisions are more likely to deal with the efficient allocation of resources on the frontier (e.g., what proportion of resources to devote to fixed assets and what proportion to remunerated labor) rather than with problems associated with the pure wastage of resources, unless of course the latter threaten the existence of the organization. If this intuition is correct, then donors would be monitoring the organization to prevent allocative rather than technical inefficiencies. In any case, a proper test of the FJW conjecture should account for both types of cost inefficiencies.

---

[4] To see why, consider the following potential bias. Suppose charities with a relatively large voluntary labor force have a larger proportion of outsider trustees to total trustees by comparison to charities with a relatively large paid labor force (a not unreasonable assumption). Since voluntary labor is costless, one would find a positive cross-sectional relationship between charity costs and the proportion of insider to outsider trustees, even if the FJW conjecture is false. Therefore, it is crucial to account for the cost effect of the input-output mix—for example, the mix between voluntary and paid labor inputs—when testing the conjecture.

We propose to test the FJW conjecture in a manner that addresses the concerns raised above. We test the conjecture that large donors monitor charities to prevent technical inefficiencies. Specifically, we measure the technical efficiency of each charity in our sample by variants of data envelopment analysis (DEA) (Charnes et al. 1978). DEA computes technical efficiency indices by comparing the organization's *ex post* input-output production levels to an industry production frontier. These indices measure the extent to which the specific organization wastes resources in comparison with the *ex post* results of other organizations in the same industry. We cross-sectionally regress these technical efficiency indices on the ratio of insider to total trustees and on the organization's debt-value ratio, where by value we mean the book value of assets. If the FJW conjecture is correct, we should find that the greater the proportion of insider to total trustees, the less efficient the organization. Thus, the regression coefficient for the ratio of insider to total trustees variable should be significant and negative.

The coefficient of the debt-value ratio should also be significant and negative because the larger the debt-value ratio, the less dependent is the institution on donor capital for financing. This in turn suggests that donors will have lesser influence on managerial activities and, ultimately, lesser ability to monitor management activities.

An alternative hypothesis is that large donors monitor the allocative efficiency of the organization. For this hypothesis we regress charity expenses per dollar of revenues, adjusted for technical inefficiencies, on the proportion of insider to total trustees and the organization's debt-value ratio. If the FJW conjecture concerning allocative inefficiencies is correct, then the coefficients of the ratio of insider to total trustees variable and the debt-value variable should be significant and positive.

III. Measuring Technical Efficiency of Charities

The unavailability of input and output price data determines to a great extent the approach one uses to measure efficiency at the organizational (firm) level. Two methodologies are proposed. In the standard econometric approach, a stochastic production or cost frontier is estimated for the industry, and each organization's efficiency is assessed relative to the estimated frontier (e.g., Kumbhakar 1988). This methodology requires knowledge of price data and, thus, is unsuitable for charities.[5]

DEA is another approach.[6] It uses maximum likelihood estimates and is statistically consistent (Banker 1992a; for surveys and other properties of DEA applications see: Banker et al. 1989a, 1989b; Banker and Thrall 1992; Borden 1988; Callen 1991; Seiford and Thrall 1990).

DEA is suited for measuring the technical efficiency of nonprofit organizations like charities because only input-output *quantity* data are required. The analysis utilizes cross-sectional, *ex post* input-output data of individual organizations in the same industry to map out the industry production function frontier. Unlike standard production theory, however, DEA does not assume a particular functional form (e.g., translog) for the industry production function, which is approximated by piecewise linear (or log-linear) functions. As shown by Banker and Maindiratta (1988), these linear approximations are determined endogenously to envelop the observed data tightly.

---

[5] Estimating the stochastic production frontier econometrically without reference to input prices yields inconsistent parameter estimates unless inputs are exogenously determined. Since, in general, inputs are endogenously determined, first-order cost minimization conditions must be appended to the production frontier or, equivalently, a stochastic cost frontier estimated. In either case, input price data are necessary.

[6] For a comparison of the econometric approach to DEA in the estimation of hospital cost functions, see Banker et al. (1986).

To avoid the limitations of assuming constant returns to scale, Banker et al. (1984) generalized the original DEA model to allow for any degree of (local) scale economies; this version of DEA is hereafter designated the BCC model.[7] Since we cannot know *a priori* if the organizations in our sample exhibit constant returns to scale, we present results for technical efficiency indices using the Charnes et al. model (1978; CCR, hereafter) and the BCC model.

Both the standard CCR and BCC indices assume that outputs and inputs are substitutable so that, for example, changes in production of one output will affect production of other outputs. However, as argued subsequently, charity outputs may not be substitutable for each other, in which case the standard DEA indices may understate the degree of organizational inefficiency. Therefore, following Banker et al. (1989b), we also estimate CCR and BCC technical efficiency indices that are adjusted for the nonsubstitutability of outputs. The appendix describes the program for obtaining DEA estimates when outputs are nonsubstitutable.

IV. The Data

Data and Variable Definitions

The *1988 Third Sector Directory Donor's Guide to Fund Raising Organizations in Canada* (hereafter *Donor's Guide*) lists about 4,500 charities that are classified into eight major categories of activities. Each of the eight major categories is further divided into subcategories, each of which is essentially an industry.[8] Because DEA yields relative efficiency measures, it is important that our analysis be confined to a specific industry. We chose the charities listed under "Specific Health Focus Organizations" (including mental health), a category that does not include facilities or medical services *per se* and that is composed of 276 charity organizations. For each charity, a copy of the latest tax information return was requested from Revenue Canada.[9]

The choice of the "Specific Health Focus" category was dictated by three considerations. First, because it is the largest charity type in the *Donor's Guide*, the choice of this industry helped to ensure a reasonable sample size that included volunteer labor data. Second, the production technology of organizations within this group is fairly homogeneous. Third, charities in this group are pure nonprofit organizations.

Of the 276 charities listed in the "Specific Health Focus" industry, we were able to obtain filed tax returns from Revenue Canada for 224 charities for the years 1986–87.[10]

---

[7] The BCC model is more useful than the CCR model because, as shown by Banker and Maindiratta (1988), an observation that is ranked inefficient by the BCC model must necessarily deviate from profit maximization.

[8] According to the *Donor's Guide*, the industry categories were obtained via an information questionnaire. The questionnaire was sent to all tax-exempt charities listed with the Charities Division of Revenue Canada. Each charity was asked to indicate its primary activity and was categorized and subcategorized on the basis of this primary activity.

[9] In Canada, tax-exempted charities must file an annual "Registered Charity Information Return and Public Information Return" form (which we refer to as the tax information return) with Revenue Canada's Taxation Department. Copies of these forms are available, for a fee, from the tax authorities.

[10] Prior to 1986, the tax form did not request volunteer labor data, making the tax returns prior to that year unsuitable for our study. Starting in 1986, this data was requested but on a voluntary basis only.

Of these 224 returns, 72 yielded data on volunteer labor. The fact that the volunteer labor data were provided on a voluntary basis raises the potential for selection bias; that is, voluntary disclosure may have been made by relatively more efficient charities. However, no specific analysis was made to determine the extent of selectivity, if any.

The tax information return includes a section describing the charity's goals and activities. A careful cross-sectional analysis yielded the output definitions. We found that the primary goal of "Specfic Health Focus" charities is to raise funds from governmental and private sources to promote activities related to a specific disease (e.g., Cancer, Alzheimer's, etc.). These activities are of two types: (1) transferring funds raised to researchers and research institutions and (2) disseminating information concerning the disease.

The above analysis led us to define organizations in this industry as primarily fund-raising institutions.[11] Outputs were defined as the total funds raised by the charity from governmental (Q_1) and private nongovernmental (Q_2) sources.[12] This distinction needs to be made because governmental sources often entail different input requirements. Government funding usually requires justification via documentation and specialized reports, which may, in turn, require specialized remunerated labor expertise. In contrast, we assert, nongovernmental funds are more readily tapped by personal connections, that is, by volunteer labor sources.[13]

The tax information return for each charity in our sample also contained a balance sheet and remuneration data for employees and trustees. Most income statement data are excluded from public purview by Revenue Canada. The available data permitted us to specify three major inputs into charitable fund-raising activities: (1) volunteer labor (L_1), measured by the number of hours (per year) contributed to the charity, (2) remunerated labor measured by total annual remunerations to all employees of the organization (L_2), and (3) capital (K), measured by the book value (original cost) of fixed assets (land, buildings, and other fixed assets).

Except for volunteer labor, all inputs are measured in dollars. This assumes that input prices are approximately constant across different charities in the "Specific Health Focus" industry.

Additionally, the tax information return includes data on the number of trustees receiving remuneration from the charity and the total amount of the remuneration. We defined trustees receiving remuneration as insider trustees, who, according to the FJW conjecture, are more likely to be concerned with their personal return (perquisites) than with running the organization efficiently. Other trustees are assumed to be interested in the charity's welfare either because they are major donors or because they are willing to

---

[11] Furthermore, charities that "are supported mainly by their respective members" and that are not engaged in major fund raising are excluded automatically from the *Donor's Guide* (1988, 133) and, therefore, did not enter the analysis.

[12] An output definition which takes into account the benefits provided to society by research activities and information dissemination would be more comprehensive. Such direct measures are unavailable, however. Using the amounts spent on research activities versus information dissemination in lieu of benefits is conceptually incorrect because expenditures are themselves a function of production efficiency.

[13] The statistically significant Pearson correlation coefficient of 0.469 (Spearman rank correlation of 0.414), between the proportion of government funding to private funding and the ratio of employee remuneration to volunteer labor, indicates that government funding, as opposed to private funding, is related to remunerated rather than volunteer labor. Thus, raising government funds may require a different technology than raising private funds, suggesting that different aspects of charity output may not be substitutable for each other.

Table 1
The Sample

Disease Type	*No. of Charities*
Neurological diseases[a]	30
Blindness	4
Cancer	5
Heart disease	4
Social welfare	9
Lung diseases	8
Mentally handicapped	6
Other[b]	6
Total	72

[a] Neurological diseases include: cerebral palsy (8), Alzheimer's (5), paraplegic (4), multiple sclerosis (4), epilepsy (2), Parkinson (2), neuro-oncology (2), and others (3).

[b] The other category includes: burns (2), drugs/alcohol (1), allergy (1), arthritis (1), and hemophilia (1).

invest time and effort, as evidenced by their willingness to be trustees without compensation. The latter group is defined as outsider trustees.

To summarize, our analysis leads us to specify two potential production technologies for charities in the "Specific Health Focus" industry: one assumes that charity outputs are substitutable and the other assumes that they are not. In the former case, the industry production function takes the form:

$$G(Q_1, Q_2, L_1, L_2, K) = 0.$$

To the nonsubstitutable case, the technology is described by two separable production functions of the form:

$$Q_1 = F_1(L_1, L_2, K),$$
$$Q_2 = F_2(L_1, L_2, K).$$

Sample Sizes

Since our total sample of 72 charities is comprised of both small and large organizations and since larger charities are more likely to incur higher agency costs, we further divided our sample into two subsamples, depending on whether total receipts (annual funds raised) are more or less than $100,000. This yielded subsamples of 47 and 25 charities for each type, respectively. In the analysis that follows, each subsample is treated as if it comprised the entire industry and the DEA indices are calculated accordingly. Table 1 provides a distribution of the sample in terms of disease types.

V. Empirical Results

Using the input-output data, we solved the DEA programs for our full sample (72 charities) and for each of the two subsamples described above (47 and 25 charities). Four indices were calculated for each organization (and for each sample): the standard CCR constant returns-to-scale index (h_0^*), the standard BCC variable returns-to-scale index (h_1^*), and the CCR and BCC indices adjusted for nonsubstitutability of outputs.

Table 2 lists the means and standard deviations of these indices and other relevant data for each sample. For comparative purposes, table 2 also includes data on the 152 charities "Specific Health Focus" category that were excluded from the analysis for lack of volunteer labor data. Table 2 indicates that the 152-charity sample is comprised of larger organizations than the 72-charity sample, which implies that the results of our analysis may not be generalizable to the entire industry.

To test the hypothesis that donors (outsider trustees) monitor charities to prevent pure resource wastage (technical inefficiencies), we cross-sectionally regressed each efficiency index on the proportion of insider trustees to total trustees and the debt-value ratio:

$$h_i^* = b_0 + b_1 PROP + b_2 DV + \epsilon, \tag{1}$$

where the b_i are parameters and ϵ is the usual white noise innovation term.

Since the DEA indices take on values between 0 and 1 only (by construction) and since some of the data took on these boundary values, the regressions could not be estimated by ordinary least squares (OLS).[14] Instead, a truncated regression maximum likelihood technique was used to estimate the coefficients.[15] In table 3, we present the regression results for the substitutable (output) technology with different DEA indices for the dependent variable. The index h_0^* is the CCR index, which presupposes constant returns to scale. The index h_1^* is the BCC index, which allows for variable returns to scale. Both indices in equation (1) yield similar results for all sample sizes.[16] The FJW conjecture should have led to significant and negative regression coefficients for *PROP* and *DV*. However, b_1 and b_2 are generally not significant and the signs of these coefficients are contrary to the hypothesized direction.[17]

It was hypothesized that the greater the ratio of insider trustees to total trustees the more inefficient the organization is likely to be. However, what may be crucial in determining technical efficiency is not necessarily the ratio of insider to total trustees but rather the power of insider trustees to appropriate more salary and perquisites for themselves.[18] One potential measure of such power is the proportion of insider trustee remuneration to total employee remuneration (*PROPA*). We replicated the regressions with *PROPA* instead of *PROP*, with similar results. Thus, we conclude that as far as technical inefficiency is concerned, the FJW conjecture that large donors mitigate against technical inefficiencies is not borne out by the data.[19]

To test the FJW conjecture in the context of allocative efficiency, we regress total costs per dollar of receipts, multiplied by the technical efficiency index, on the ratio of

---

[14] It is well-known that with limited dependent variables, OLS may severely bias the coefficient estimates (Maddala 1983).

[15] The nonlinear optimization was solved by the Newton-Raphson method modified by the Davidson-Fletcher-Powell iteration as described in Quandt (1983). The truncated regression estimation assumes that the limited dependent variable is truncated normal.

[16] Our empirical results are not sensitive to the particular DEA efficiency index h_i^*, $i=0,1$. This is not surprising perhaps since the indices are highly correlated. The product moment correlation between h_0^* and h_1^* for the 72-charity sample, for example, is 0.71 for the substitutable technology and 0.72 for the nonsubstitutable technology.

[17] OLS yielded similarly qualitative results.

[18] The United Way of America scandal is supportive of this argument.

[19] In case technical efficiency is a function of the age of the institution, with older institutions perhaps less efficient, we reran the regressions in table 3 with age as an additional independent variable. This variable turned out to be insignificant.

Table 2
Summary Statistics

	Sample Size							
	Full Sample With Labor Data ($n=72$)		Subsample With Total Receipts of at Least \$100,000 ($n=47$)		Sample With Total Receipts<\$100,000 ($n=25$)		Excluded Charities With No Labor Data Reported ($n=152$)	
Attributes	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*
TA	744,996	1,490,477	1,106,936	1,727,199	36,196	52,694	1,370,386	3,406,983
TR	809,820	1,426,918	1,215,481	1,630,307	47,179	31,817	1,905,244	4,691,299
TD	811,528	1,516,374	1,215,821	1,751,313	51,455	43,330	1,632,338	3,990,135
TC	171,718	255,286	251,877	284,880	21,021	31,750	323,614	717,203
PROP	0.144	0.184	0.176	0.203	0.084	0.125	0.124	0.219
PROPA	0.188	0.267	0.193	0.221	0.180	0.343	0.143	0.236
DV	0.433	0.639	0.513	0.732	0.311	0.389	0.306	0.345
RC	0.286	0.287	0.269	0.214	0.318	0.393	0.311	0.377
h_0^*	0.478	0.320	0.605	0.291	0.443	0.362	N.A.	N.A.
h_1^*	0.645	0.321	0.728	0.280	0.745	0.341	N.A.	N.A.
nh_0^*	0.387	0.303	0.518	0.309	0.361	0.327	N.A.	N.A.
nh_1^*	0.541	0.327	0.634	0.304	0.653	0.339	N.A.	N.A.
Q_1	113,576	475,447	171,168	582,245	5,304	16,013	220,787	638,190
Q_2	505,161	1,025,244	760,508	1,195,867	25,109	28,611	1,439,804	4,251,294
L_1	27,064	92,264	40,228	112,369	2,315	2,155	N.A.	N.A.
L_2	219,402	325,573	324,521	361,649	21,779	23,981	442,907	1,423,514
K	182,284	495,731	273,566	594,879	10,674	41,739	429,282	924,576

Notes: *TA*=total assets; *TR*=total receipts; *TD*=total disbursements; *TC*=total costs; *PROP*=proportion of insider trustees to total trustees; *PROPA*=proportion of insider trustee remuneration to total employee remuneration; *DV*=debt-value ratio (debt to total assets); *RC*=proportion of total costs (*TC*) to total receipts (*TR*); h_i^*=technical efficiency index $i=0, 1$ for substitutable technologies; nh_i^*=technical efficiency index $i=0,1$ for nonsubstitutable technologies; Q_1=funds raised from governmental sources; Q_2=funds raised from private sources; L_1=volunteer labor; L_2=remunerated labor; K=capital (buildings and other fixed assets); and N.A.=not available.

Table 3
Technical Efficiencies
Limited Dependent Variable Regression Results for Substitutable Technologies
Log-likelihood, Coefficient Estimates, and *t*-ratios

	Samples		
Model	*72–Charity Sample*	*47–Charity Subsample*	*25–Charity Subsample*
$h_i^* = b_0 + b_1 PROP + b_2 DV + \epsilon$			
CCR index:			
h_0^* (*LL*)	5.037	14.716	7.797
Constant	0.164	0.359	0.152
	(0.665)	(6.886)	(0.698)
PROP	0.278	0.194	−0.622
	(0.544)	(1.133)	(−0.537)
DV	0.054	0.151	0.008
	(0.258)	(3.167)	(0.024)
BCC index:			
h_1^* (*LL*)	3.480	4.685	4.902
Constant	0.344	0.510	0.271
	(2.482)	(5.419)	(2.091)
PROP	0.142	0.032	−0.253
	(0.417)	(0.092)	(−0.351)
DV	0.254	0.148	0.517
	(1.412)	(0.889)	(2.858)
$h_i^* = b_0 + b_1 PROPA + b_2 DV + \epsilon$			
CCR index:			
h_0^* (*LL*)	4.996	13.907	7.510
Constant	0.163	0.391	0.063
	(0.663)	(8.334)	(0.201)
PROPA	0.181	−0.032	0.050
	(0.420)	(−0.169)	(0.065)
DV	0.076	0.166	0.055
	(0.363)	(3.047)	(0.148)
BCC index:			
h_1^* (*LL*)	3.400	6.285	4.800
Constant	0.362	0.452	0.225
	(2.610)	(5.664)	(1.364)
PROPA	0.011	0.442	0.033
	(0.032)	(1.699)	(0.081)
DV	0.259	0.087	0.564
	(1.398)	(0.534)	(2.629)

Note: The *t*-statistics are in parentheses. The variable definitions are: h_i^*=DEA efficiency index $i=0,1$; *LL*=log-likelihood; *PROP*=proportion of insider trustees to total trustees; *PROPA*=proportion of insider trustee remuneration to total employee remuneration; and *DV*=debt-value ratio (debt to total assets).

insider trustees to total trustees (as suggested by Williamson 1983) and the debt-value ratio. Following Banker and Maindiratta (1988), we define allocative efficiency to be aggregate organizational efficiency (i.e., actual costs divided by minimum costs) divided by technical efficiency.[20] Therefore, we multiply costs per dollar of receipts by the technical efficiency index so that any remaining inefficiencies are allocative in nature. More formally, the regression equation is:

$$RC \cdot h_i^* = a_0 + a_1 PROP + a_2 DV + \epsilon, \tag{2}$$

where the a_k are estimated parameters.[21] If large donors mitigate against allocative inefficiencies, we expect $a_1 > 0$ and $a_2 > 0$.

Table 4 presents the results of regression equation (2). Both indices (h_i^*, $i = 0, 1$) yield essentially the same results. The signs of the *DV* coefficient are in the hypothesized direction and are generally significant. Nevertheless, the sign of the *PROP* variable is often contrary to what was hypothesized, and its coefficient, with one exception, is generally insignificant. As before, we replace *PROP* with *PROPA*, but we obtain similar results.[22]

As table 4 shows, costs per dollar of receipts are almost always significantly and positively related to the organization's debt-value ratio. This result, however, may simply reflect the interest expense component of costs and may not be a confirmation of the FJW hypothesis. Unfortunately, the publicly accessible tax return data did not include explicit interest expense, or most income statement data. We, therefore, tried to obtain the interest expense data from financial statements which were either publicly available or solicited from the charities. In total, these efforts yielded the relevant interest expense data for 39 of the 72 charities in our sample.

Table 4 also shows regression results for the 39-charity subsample with interest expense deducted from charity costs (the dependent variable). The results do not show any significant *PROP* (or *PROPA*) coefficient. The debt-value ratio is also no longer significant, suggesting that the positive significant relationship between costs (per dollar of revenues) and the debt-value ratio may be due to interest expense.

The empirical analysis in tables 3 and 4 assumed a substitutable technology. Repeated analysis under the assumption of no output substitution, however, showed that the results are robust to this assumption.

---

[20] Banker and Maindiratta also developed a procedure for decomposing aggregate efficiency into separate technical and allocative components that yield upper- and lower-bound DEA estimates. This procedure maintains the assumption of profit maximization. These bounds are not relevant here because the behavioral assumption of profit maximization is not appropriate for charities. In any case, these bounds cannot be computed because input price data are unavailable.

[21] Given the assumptions of cross-sectionally constant prices and technology, the dependent variable in equation (2) is an exact measure of allocative efficiency provided the production technology exhibits constant returns to scale ($RC \cdot h_0^*$). With variable returns to scale, the dependent variable ($RC \cdot h_1^*$) is an inexact measure of allocative efficiency. However, there is no particular direction to the noise, which is assumed to be captured by the error term in the regression. We also estimated equation (2) in the more general log-linear form:

$$\ell n\, RC = C_0 + C_1 \ell n\, PROP + C_2 \ell n\, DV + C_3 \ell n\, h_i^* + \epsilon.$$

As is to be expected, the C_3 coefficient was generally significant and negative. However, because the C_1 and C_2 coefficients exhibited properties very similar to the results in table 4, we do not report the results of the log-linear regressions.

[22] We also ran regression (2) with the output mix and input mixes as independent variables. Since the results are not qualitatively different from those in table 4, we do not report them.

Table 4
Allocative Efficiencies
OLS Regression Results for Substitutable Technologies
Coefficient Estimates and *t*-ratios

	Samples			
Model	72–Charity Sample	47–Charity Subsample	25–Charity Subsample	39–Charity Subsample
$RC \cdot h_i^* = a_0 + a_1 PROP + a_2 DV + \epsilon$				
$RC \cdot h_0^*$				
Constant	0.091 (3.600)	0.113 (3.635)	0.110 (2.261)	0.113 (4.202)
PROP	−0.072 (−0.735)	0.005 (0.048)	−0.019 (−0.072)	0.049 (0.532)
DV	0.120 (4.223)	0.110 (3.721)	0.094 (1.115)	−0.007 (−0.142)
F	8.94	7.08	0.66	0.922
R^2	0.206	0.243	0.057	−0.008
$\bar{R}^2$	0.183	0.209	−0.029	−0.047
$RC \cdot h_1^*$				
Constant	0.060 (2.521)	0.145 (3.947)	0.091 (0.886)	0.146 (4.215)
PROP	0.128 (1.921)	−0.020 (−0.156)	0.174 (0.315)	0.054 (0.460)
DV	0.110 (3.973)	0.120 (3.406)	0.476 (2.687)	0.010 (0.156)
F	10.91	5.85	3.62	0.873
R^2	0.240	0.210	0.248	0.008
$\bar{R}^2$	0.218	0.174	0.179	0.048

(Continued)

It is possible that the disease subgroups in table 1 are not sufficiently homogeneous to constitute an industry. We therefore repeated the regression analysis (eqs. [1] and [2]) using an intercept dummy variable for each subclass. The results were qualitatively similar to those obtained previously.

In a further attempt to address this issue, we reestimated our technical efficiency indices for the 30 organizations in the neurological diseases category—the largest subclass listed in table 1—thereby treating this subclass as a separate industry. Again, the results of estimating equations (1) and (2) were qualitatively similar to those obtained for the other samples.

VI. Conclusion

Following a conjecture by FJW, we hypothesized that large donors to charity organizations mitigate against technical or allocative inefficiencies. Testing these hypoth-

Table 4—*Continued*

	Samples			
Model	72–Charity Sample	47–Charity Subsample	25–Charity Subsample	39–Charity Subsample
$RC \cdot h_i^* = a_0 + a_1 PROPA + a_2 DV + \epsilon$				
$RC \cdot h_0^*$				
Constant	0.060	0.097	0.092	0.125
	(2.521)	(3.224)	(2.045)	(4.431)
PROPA	0.128	0.096	0.073	−0.037
	(1.921)	(0.957)	(0.777)	(−0.509)
DV	0.110	0.104	0.104	−0.003
	(3.973)	(3.440)	(1.260)	(−0.059)
F	10.91	7.81	1.09	0.130
R^2	0.240	0.262	0.90	0.007
$\bar{R}^2$	0.218	0.228	0.008	−0.048
$RC \cdot h_1^*$				
Constant	0.092	0.125	0.074	0.164
	(3.573)	(3.499)	(0.781)	(4.546)
PROPA	0.113	0.116	0.156	−0.068
	(1.585)	(0.119)	(0.790)	(−0.737)
DV	0.124	0.110	0.486	0.014
	(4.140)	(3.069)	(2.790)	(0.225)
F	10.86	6.43	3.97	0.741
R^2	0.239	0.226	0.265	0.017
$\bar{R}^2$	0.217	0.191	0.199	−0.038

Note: The t-statistics are in parentheses. The variable definitions are: RC = proportion of total costs to total receipts; $PROPA$ = proportion of insider trustee remuneration to total employee remuneration; DV = debt-value ratio (debt to total assets); and h_i^* = DEA efficiency index $i=0,1$.

eses on a sample of Canadian "Specific Health Focus" charities, the results, by and large, did not confirm these hypotheses.

Our sample was limited to charities for which volunteer labor data were available. Those charities were found to be smaller than those of the rest of the "Specific Health Focus" charity industry. Thus, whether the FJW conjecture can be rejected for charities in other industries, or even for all charities in the "Specific Health Focus" industry, is an open question. Further, we utilized specific versions of DEA. While this analytical tool is particularly suited for this study's task, other methods might yield different results. Future researchers may examine these issues.

Because the number of major donors in each charity is not directly observable, we used proxy variables such as the number of outsider trustees or the debt-to-value ratio. This problem however, is likely to plague further studies, which rely on publicly available data, as well.

Independently of the FJW conjecture, our empirical results imply that neither the composition of the board of trustees nor the debt-value ratio affects the technical or allocative efficiency of nonprofit organizations. In particular, even if outsider trustees donate time or technical expertise rather than money, the proportion of outsiders on the board does not appear to have any effect on organizational efficiency.

Appendix

DEA Technical Efficiency Indices for Nonsubstitutable Technologies

One crucial assumption underlying both the standard CCR and BCC DEA indices is that outputs (and inputs) are substitutable for each other. As emphasized in the text, however, charity outputs may not be substitutable for each other, which means that standard technical efficiency indices overstate the degree of organizational inefficiency. Following Banker et al. (1989b), who analyze the case of nonsubstitutable input technologies, we restated the standard DEA programs to allow for nonsubstitutable output technologies. Formally, let X_{ij}, $Y_{rj} > 0$ represent input-output data for charity j, $j = 1, \ldots, n$ where $i = 1, \ldots, m$ is the index of the inputs, and $r = 1, \ldots, s$ is the index of the outputs. The subscript o in the objective function denotes one of the n charities, specifically, the one whose efficiency is currently being measured. For each output r (and organization o), we solve the extremal problem:

$$\begin{array}{lll} \underset{h_{ro}, h_{rj}}{\text{Minimize}} & h_{ro} & \\ \text{subject to:} & -\sum_{j=1}^{n} h_{rj} X_{ij} + h_{ro} X_{io} \geq 0 & i = 1, \ldots, m;\ r = 1, \ldots, s; \\ & \sum_{j=1}^{n} h_{rj} Y_{rj} \geq Y_{ro} & r = 1, \ldots, s; \\ & \sum_{j=1}^{n} h_{rj} = 1 & r = 1, \ldots, s; \text{ and} \\ & h_{rj} \geq 0 & r = 1, \ldots, s;\ j = 1, \ldots, n. \end{array}$$

This yields s output efficiency indices for organization o. The aggregate technical efficiency index for this organization is computed as a weighted average of its output efficiency indices, with the outputs as weights, that is:

$$\left(\sum_{r=1}^{s} h^{*}_{ro} Y_{ro}\right) \Big/ \left(\sum_{r=1}^{s} Y_{ro}\right).$$

This aggregate technical efficiency measure is a BCC technical efficiency index.

The CCR aggregate technical efficiency index for nonsubstitutable output technologies is solved in a similar fashion after deleting constraints of the form:

$$\sum_{j=1}^{n} h_{rj} = 1$$

from the extremal problem.

References

Babcock, C. R. 1992. 3 locals ask national United Way about president's spending. *Washington Post* (February 20): A6.

Banker, R. D. 1992a. Maximum likelihood, consistency, and data envelopment analysis. *Management Science* 38 (Forthcoming).

——. 1992b. Selection of efficiency evaluation models. *Contemporary Accounting Research* 9 (Forthcoming).

——, A. Charnes, and W. W. Cooper. 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30 (September): 1078–92.

——, ——, ——, J. Swarts, and D. Thomas. 1989a. An introduction to data envelopment analysis with some of its models and their uses. *Research in Governmental and Nonprofit Accounting* 5: 25–64.

——, R. F. Conrad, and R. P. Strauss. 1986. A comparative application of data envelopment analysis and translog methods: An illustrative study of hospital production. *Management Science* 32 (January): 30–44.

——, S. Das, and S. M. Datar. 1989b. Analysis of cost variances for management control in hospitals. *Research in Governmental and Nonprofit Accounting* 5: 269–91.

——, and A. J. Maindiratta. 1988. Nonparametric analysis of technical and allocative efficiencies in production. *Econometrica* 56 (November): 1315–32.

——, and R. M. Thrall. 1992. Estimation of returns to scale using data envelopment analysis. *European Journal of Operational Research* 56 (Forthcoming).

Borden, J. P. 1988. An assessment of the impact of diagnoses—related group (DRG)—based reimbursements on the technical efficiency of New Jersey hospitals using data envelopment analysis. *Journal of Accounting and Public Policy* 7 (Summer): 77–96.

Callen, J. L. 1991. Data envelopment analysis: Partial survey and applications for management accounting. *Journal of Management Accounting Research* 3: 35–56.

Charnes, A., W. W. Cooper, and E. Rhodes. 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2 (November): 429–44.

Clotfelter, C. T., and L. M. Salamon. 1982. The impact of the 1981 Tax Act on individual charitable giving. *National Tax Journal* 35 (June): 171–87.

Easley, D., and M. O'Hara. 1986. Optimal nonprofit firms. In *The Economics of Nonprofit Institutions*, edited by S. Rose-Ackerman. New York: Oxford University Press, 85–93.

Falk, H. 1992. Towards a framework for not-for-profit accounting. *Contemporary Accounting Research* 8 (Spring): 468–99.

Fama, E. F., and M. C. Jensen. 1983a. Separation of ownership and control. *Journal of Law and Economics* 26 (June): 301–25.

——, and ——. 1983b. Agency problems and residual claims. *Journal of Law and Economics* 26 (November): 327–49.

Farrell, M. J. 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society* A 120: 253–81.

Feldstein, M. S., and C. T. Clotfelter. 1976. Tax incentives and charitable contributions in the United States. *Journal of Public Economics* 5 (January/February): 1–26.

Hansman, H. B. 1980. The role of nonprofit enterprise. *The Yale Law Journal* 8 (April): 835–98.

Hochman, H. M., and J. D. Rodgers. 1977. The optimal tax treatment of charitable contributions. *National Tax Journal* 30 (March): 1–19.

Kumbhakar, S. C. 1988. On the estimation of technical and allocative inefficiency using stochastic frontier functions: The case of U.S. class 1 railroads. *International Economic Review* 29 (November): 727–43.

Maddala, G. S. 1983. *Limited and Qualitative Dependent Variables in Economics.* New York: Cambridge University Press.

Miller, A., A. Wolfberg, and B. Shenitz. 1992. At United Way charity began at home. *Newsweek* (March 9): 56.

Niskanen, W. A. 1971. *Bureaucracy and Representative Government.* Chicago: Aldine.

Pauly, M. 1980. Delivered comments. In *The Economics of Nonproprietary Organizations*, Supplement 1 to *Research in Law and Economics.* Greenwich, CT: JAI Press, 170–75.

Quandt, R. E. 1983. Computational problems and methods. In *Handbook of Econometrics*, Vol. 1, edited by Z. Griliches and M. D. Intrilligator. Amsterdam: North-Holland, 699–764.

Rose-Ackerman, S. 1982. Charitable giving and "excessive" fundraising. *The Quarterly Journal of Economics* 97 (May): 195–212.

Seiford, L. M, and R. M. Thrall. 1990. Recent developments in DEA: The mathematical programming approach to frontier analysis. *Journal of Econometrics* 46 (December): 7–38.

Shepard, C. E. 1992. Perks, privileges and power in the nonprofit world. *Washington Post* (February 16): A1, A38.

Steinber, R. 1986. Should donors care about fundraising? In *The Economics of Nonprofit Institutions*, edited by S. Rose-Ackerman. New York: Oxford University Press, 347–64.

Strnat, J. 1986. The charitable contribution deduction: A politico-economic analysis. In *The Economics of Nonprofit Institutions*, edited by S. Rose-Ackerman. New York: Oxford University Press, 265–96.

The Third Sector Directory *Donor's Guide to Fund Raising*. 1988. Toronto, Ontario, Canada: Third Sector Publishing.

Thompson, E. A. 1980. Charity and nonprofit organizations. In *The Economics of Nonproprietary Organizations*, Supplement 1 to *Research in Law and Economics*. Greenwich, CT: JAI Press, 125–38.

Weisbrod, B. A. 1975. Toward a theory of the voluntary nonprofit sector in a three-sector economy. In *Altruism, Morality and Economic Theory*, edited by E. S. Phelps. New York: Russell Sage Foundation, 171–95.

———. 1980. Private goods, collective goods: The role of the non-profit sector. In *The Economics of Nonproprietary Organizations*, Supplement 1 to *Research in Law and Economics*. Greenwich, CT: JAI Press, 139–77.

———, and M. Schlesinger. 1986. Public, private, nonprofit ownership and the response to asymmetric information: The case of nursing homes. In *The Economics of Nonprofit Institutions*, edited by S. Rose-Ackerman. New York: Oxford University Press, 132–51.

Whinston, A. 1980. Discussion. *The Economics of Nonproprietary Organizations*, Supplement 1 to *Research in Law and Economics*. Greenwich, CT: JAI Press: 184–85.

Williamson, O. E. 1983. Organization form, residual claimants, and corporate control. *Journal of Law and Economics* 26 (June): 351–66.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 66–88

Measuring Production Efficiency in a Not-for-Profit Setting: An Extension

Yaw M. Mensah
Shu-Hsing Li
Rutgers University

SYNOPSIS AND INTRODUCTION: In a recent study, Hayes and Millar (1990) presented empirical evidence on the cost function and apparent cost-optimizing behavior of the local managers of 33 county jails in Tennessee. Two main arguments were advanced: (1) line-item budgeting (LIB) was an ineffective control because the cost shares are assumed fixed in such budgetary settings, and (2) the budgetary control and performance evaluation process in not-for-profit settings could be improved if the underlying cost function was estimated by using a flexible functional form such as the translog to gain knowledge of possible input substitution and output transformation. This translog budget model was viewed implicitly as a superior alternative to the more traditional LIB approach.

The main objective of this study is to extend Hayes and Millar's idea of a translog budget model by outlining a more complete budgetary system. It is shown that a frontier cost function generated from the ordinary least squares (OLS) translog function can be used to identify four types of inefficiencies: the degree of technical, allocative, and scale inefficiency as well as institutional X-inefficiency (Liebenstein 1966). These four types of inefficiencies are then linked to both the long-term and short-term objectives of budgeting; namely, performance evaluation, subordinate manager motivation, planning, and control.

The second objective of the study is to compare the translog budget model (as revised) against the most frequently applied alternative technique, the nonparametric data envelopment analysis (DEA), in the same context. From the standpoint of a routine budgetary control system, it was found that the DEA model is more suitable than the translog model.

This article has benefited from the comments of participants at accounting seminars at Rutgers University and New York University. We are particularly grateful for the comments of Ajay Maindiratta, Bin Srinidhi, Bikki Jaggi, Ephraim Sudit, Ted Hollander, and the two anonymous referees of this journal.

Submitted October 1991.
Accepted June 1992.

In not-for-profit settings, however, an econometric model such as the translog may be needed initially in the specification of the most appropriate input and output measures. Although the effectiveness of LIB as a control is not addressed directly, our analysis shows that the line-item data mandatory with this approach is an essential part of any effective budgetary system designed for a not-for-profit setting.

Key Words: *Data envelopment analysis, Line-item budgeting, Production efficiency evaluation, X-efficiency.*

Data Availability: *The data used in this study are included as an appendix.*

THE remainder of the article is organized as follows. In section I we discuss our primary reservations about Hayes and Millar's principal argument against LIB. We describe our extension of their translog budget model and our version of a DEA budget model in section II. In section III we present an illustration of the alternative models using the same data as Hayes and Millar (1990). The respective merits of the alternative models are discussed in section IV, and our conclusions are presented in the final section.

I. A Perspective on the Hayes-Millar Study

Cross-Sectional Constancy of Cost Shares

Common in not-for-profit settings (particularly in governmental environments), LIB achieves financial control through limits on expenditures by type of input (the line items) rather than by the individual programs (or outputs) of the local operating units. In essence, the goal is to achieve financial control by focusing attention on the use of inputs (Ramanathan 1982, 267–68). The practice appears to be the dominant budgeting procedure among not-for-profit organizations (Lynn and Freeman 1983, 60–75).

In their study of cross-sectional data for 33 county jails in Tennessee, Hayes and Millar criticized LIB on the grounds that it assumes that the cost shares are fixed. Using a translog model as a second-order approximation of the underlying unknown cost function, they found that the cost-minimizing input cost shares were not fixed but responsive to cross-sectional differences in relative input prices and output. They interpreted this finding as an indication that LIB is deficient as a control since it imposes constraints that do not motivate managers to deviate from the cost standards implied by the line-item budget. This could result in suboptimal results, compared to approaches that allow individual local managers to alter the input proportions in response to changes in relative prices. According to Hayes and Millar (1990, 513), "line-item budgeting as a control/evaluation device is deficient. This deficiency exists because, for line-item budgets to be ideal standards, the underlying production function must be characterized by fixed cost shares."

Note, however, that this conclusion relates to the constancy of the optimal cost shares in a cross-sectional context. LIB does require that the optimal cost shares be fixed for the duration of the budget period for a given local operating unit, but nothing in the line-item theory or practice LIB suggests that the cost shares be cross-sectionally constant. That is, if for local unit A, 30 percent of the budget is assigned to personnel costs and 70 percent to other costs, it does not follow that the exact same proportions must be imposed on local unit B. The assumption of cross-sectionally fixed cost shares (i.e., cost shares whose optimality is independent of the relative prices of the inputs) requires a particularly restrictive cost/production function, specifically, the Cobb-Douglas function (Darrough 1988; Schmidt and Lovell 1979).

The potential deficiency in LIB lies in the fact that, in an environment of changing relative input prices, it may inhibit the discretionary power of the local unit managers to substitute one input for another. The issue is thus solely one of choosing the right mix of inputs, given their relative prices. That is, in the terminology introduced by Farrell (1957), allocative efficiency is required, not technical efficiency (which may involve sheer waste in the use of the inputs) or total productive efficiency (the product of the two components).[1] These concepts have been previously explored in the accounting literature, although heretofore in manufacturing settings and for production functions rather than cost functions (Barlev and Callen 1986; Callen 1988; Darrough 1988; Marcinko and Petri 1984; Mensah 1982).

However, the issue is not the type of underlying production (or cost) function, but whether LIB necessarily leads to allocative inefficiency. The existence of such inefficiency is independent of the functional form of the cost function (Bauer 1990; Forsund et al. 1980; Greene 1980a; Kopp 1981; Kopp and Diewert 1982).[2]

Completeness of Translog Budget Model

Hayes and Millar (1990, 518) proposed the translog cost function as the basis for a "budget (i.e., cost share) approach . . . capable of providing evidence of cost-efficient behavior, and . . . therefore a useful performance evaluation tool." Our concern here is that the cost share budget approach as proposed is incomplete and may in fact lead to the perception that all local units are cost-efficient even when they are not. Specifically, the proposed translog model can be extended to provide information that furthers the four principal objectives of budgeting: (1) motivation of subordinates, (2) provision of a benchmark for evaluating *ex post* performance, (3) planning, and (4) control.

Our proposed extension of Hayes and Millar's translog budget model focuses on (1) the derivation of a "best practice" frontier function from the OLS "average practice" function derived by Hayes and Millar, (2) the decomposition of the residuals from the frontier function into components of technical and allocative efficiency, (3) the derivation of information relating to economies of scale, and (4) the derivation of information relating to possible institutional X-inefficiency. Since the DEA model presents a feasible alternative to the translog in all these areas, we examine the DEA performance in these situations as well.

For a budget to be an effective short-term control device, it must provide feedback to local managers not only about *ex post* deficiencies in performance but also about the possible causes and necessary corrective action. The decomposition of the total deviation of actual costs from minimum feasible costs permits the sources of the deviation to be traced. With LIB, local managers cannot be held responsible for allocative inefficiencies since the system effectively insulates them from such responsibility. However, managers are and should be held responsible for any technical inefficiencies. In a non-LIB context, the decomposition is still helpful since it motivates local managers to respond to changes in the relative input price structure, as well as to eliminate any sheer inefficiency in the use of inputs.

---

[1] For an analysis and explicit modeling of the conditions under which LIB is costless relative to an alternative such as program budgeting, see Mensah (1992a).

[2] From this perspective, the tests performed by Hayes and Millar (1990) can be interpreted as joint tests for the appropriate functional form for the cost function. The usual symmetry of the cost function parameters and the associated cost share parameters implies that if all the tests had indicated jointly insignificant price-to-price and price-to-output interaction terms, the underlying cost function would have been a homogeneous Cobb-Douglas function.

As demonstrated later, the computation and decomposition of the total productive efficiency can be derived by the central administrator from the cost function if the local units report total costs by input type, input quantities used, as well as total output levels achieved. Since this level of disaggregation is required under LIB, it allows a central administrator to evaluate performance using either the translog or the DEA approach. Local managers are likely to be highly motivated if they are aware that the central administrator has a rigorous tool available to evaluate comparative performance and to impose post-implementation rewards or penalties. Thus, a necessary part of any budgetary control system is the ability to evaluate cross-sectional differences in performance (or, for a single unit, differences in performance over time).

Moreover, both the translog and the DEA techniques can be applied in the initial stages of budget preparation to ferret out budgetary slack. As long as there is no *a priori* collusion between local managers, the techniques discussed here can be applied to budgets proposed by the individual units to determine the degree of both allocative and technical inefficiency. Steps can then be taken to remove slack or at least obtain an adequate explanation from the local managers. Such explanations may involve unsuspected limitations in the translog or DEA models relating to input substitution possibilities. Used in this preventive mode, the techniques to be discussed can be conceived of as providing a flexible budget. Thus, a direct link to the standard costing and variance analysis approach advocated in the literature is evident (see Callen 1988; Darrough 1988).

From a long-term perspective, both the translog and DEA models can improve the planning role of budgets. A post-implementation evaluation of performance focused on cross-sectional differences in output unit costs must deal with economies of scale. A central administrator needs to know if local units of different sizes operate at different points on the long-run cost curve, since some portion of the differences in output unit costs may be due to scale inefficiency. Such knowledge is obviously useful in capital-budgeting decisions that require specification of the size of a future unit. Fortunately, economies of scale are relatively easy to extract with both the translog and DEA models.

Finally, we need to consider the possible applicability of Liebenstein's X-efficiency theory in not-for-profit settings. Liebenstein (1966, 1976) has argued that the focus on technical, allocative, and scale inefficiencies is often misplaced; frequently, the largest possible savings are achieved through questioning basic institutional assumptions. In other words, the largest potential savings can often be realized by a fundamental reorganization of the institutional structure, not by a focus on the three other sources of inefficiency already identified.[3]

---

[3] For an overview of the X-efficiency theory and empirical evidence supporting it, see Liebenstein (1976).

Cost functions rather than production functions are usually the focus of an econometric model of an industry when two assumptions are met: (1) the output of the local units can be assumed to be exogenous, as for example in regulated industries or for most not-for-profit organizations; and (2) input prices can be assumed to be exogenous (Greene 1980b; Kumbkahar 1988). The basic assumption that needs to be questioned in not-for-profit settings is the rationale for the existence of cross-sectional differences in input prices.

In for-profit settings, the competitive threat to survival forces firms with high unit costs to consider adopting competitors' production methods and ideas in order to survive. Since not-for-profit units rely on voluntary donations or subsidies from government authorities (or both), the threat of insolvency from purely competitive pressures may be perceived as remote (particularly for governmental agencies). Thus, this effective market mechanism may not work properly in such settings. Absent such an external mechanism, an internal system must be installed as part of an effective budgetary system to evaluate the effects of institutional inefficiency. We propose such a system through the construction of what we have called an index of input price advantage.

II. The Translog and DEA Budget Models

Translog Budget Model

Although there are quite a few alternative approaches to estimating a "best-practice" (or full-frontier) cost function with an econometric approach, it is well-established that OLS estimates of the slope parameters of a given functional form are consistent, although the intercept term is biased. Consequently, a full-frontier cost function can be derived as a neutrally displaced OLS function by using the method of moments. Essentially, this involves shifting the intercept of the OLS function until at least one of the residuals (the support point) is zero and all the others are positive (Bauer 1990, 54; Forsund et al. 1980, 12; Greene 1980a, 34). Once the frontier function is derived, an approach outlined in Kopp and Diewert (1982) or a simplification by Zieschang (1983) can be used to decompose the overall residual (the total productive efficiency measure) into the technical and allocative efficiencies for any unit.[4]

In conformance with common practice in econometric studies, we estimated jointly the translog cost function and two cost share equations using Zellner's seemingly unrelated regressions approach. The exact equations estimated for the Tennessee county jail data (Hayes and Millar 1990) can be written as follows for total cost (*TC*) and cost shares (*CS*):

$$TC_s^a = A_O + \sum_i^n \gamma_i \ell n(p_{si}) + \sum_k^m \phi_k \ell n(y_{sk}) + \frac{1}{2}\sum_i^n \sum_j^n \alpha_{ij} \ell n(p_{si}) \ell n(p_{si}) \ell n(p_{sj})$$

$$+ \sum_i^n \sum_k^m \beta_{ik} \ell n(p_{si}) \ell n(y_{sk}) + \frac{1}{2}\sum_k^m \sum_q^m \omega_{kq} \ell n(y_{sk}) \ell n(y_{sq}) + \epsilon_s. \tag{1a}$$

---

[4] The approach used in this article results in what is known as a deterministic cost frontier. Recent interest in the cost/production frontier literature has focused on stochastic frontiers. The problem with stochastic frontiers in this context is that the allocative efficiency ratings for individual units cannot be identified; one can only assess the mean allocative inefficiency of all the units as a group. The adoption of a deterministic frontier approach avoids this problem. For a discussion of this issue, see Bauer (1990), Ferrier and Lovell (1990), and Forsund et al. (1980).

$$CS^a_{s1}=\gamma_1+\sum_j^n \alpha_{1j}\ell n(p_{sj})+\sum_k^m \beta_{1k}\ell n(y_{sk})+\epsilon_t, \tag{1b}$$

$$CS^a_{s2}=\gamma_2+\sum_j^n \alpha_{2j}\ell n(p_{sj})+\sum_k^m \beta_{2k}\ell n(y_{sk})+\eta_s, \tag{1c}$$

where:

p_{si}=unit prices for labor (LP), boarding (BP), and maintenance (MP), respectively for unit s;
y_{sk}=jail days (JD), average term (AT), and technology (TE), respectively;[5] and
CS^a_{si}=actual cost share of input i for unit s.

To correspond to a well-behaved production function, the translog function must satisfy, *inter alia*, the theoretical requirement of linear homogeneity of degree one in input prices, that is,

$$\sum_{i=1}^{3}\gamma_i=1;\quad \sum_{i=1}^{3}\alpha_{ij}=0;\quad \text{and}\quad \sum_{i=1}^{3}\beta_{ik}=0 \qquad \text{for all } i,j \text{ inputs and } k \text{ outputs.}$$

Kopp and Diewert's (1982) decomposition technique essentially derives total costs (TC^b_s) for any local unit s at a point on the cost frontier that maintains the same input mix as currently observed. Thus, TC^b_s represents costs for a technically efficient production regime, although one that is possibly allocatively inefficient, given the actual input prices for that unit. Subsequently, costs at the technically and allocatively efficient point (TC^m_s) can be derived by evaluating the frontier cost function, given the current input prices. From Kopp (1981), a generalized measure of the Farrell (1957) input-based technical efficiency measure (λ^T) and the allocative efficiency measure (λ^A) can be derived as follows:[6]

$$\lambda^T_s=TC^b_s/TC^a_s, \tag{2a}$$

$$\lambda^A_s=TC^m_s/TC^b_s, \text{ and} \tag{2b}$$

$$\lambda^P_s=TC^m_s/TC^a_s, \tag{2c}$$

---

[5] The definition of outputs has been changed slightly from that in Hayes and Millar (1990). Average term (AT), defined as total jail days divided by the number of admissions, has been added as an extra output variable, for the same reason that Caves et al. (1981) included both passenger miles (ton miles of freight) and the average length of passenger trips (freight haul) as output measures in a study of the variable cost structure of railroads. In the county jail context, it is plausible that some part of the total costs listed in the appendix will be more dependent on the number of admissions than on total time spent in the jail. Thus, AT is expected to be inversely correlated with total operating costs, and this effect will be in addition to that of total jail days on total costs. The proxy for the capital-labor ratio (i.e., differences in technology, TE) has been redefined as the ratio of the average number of occupied beds to the number of employees. This measure ensures that differences in capacity utilization are not confounded with differences in technology.

[6] For any given frontier cost function, the Kopp-Diewert decomposition approach involves solving a system of $2n-1$ nonlinear equations (for each s unit) of the form:

$$x^b_{si}=C(y,p^b_j,1)/p^b_{si}{}^*CS^a_{si}, \qquad i=1,\ldots,n \text{ inputs}$$

$$x^b_{s1}/x^b_{sj}=x^a_{s1}/x^a_{sj}, \qquad j=2,\ldots,n \text{ inputs}$$

where x^b_{si}=vector of input quantities at the technically efficient, but allocatively inefficient, point (denoted as point B); x^a_{si}=vector of actual input quantities used; p^b_{si}=vector of unknown prices at which the current mix of input quantities would have been both technically and allocatively efficient; and $C(y,p^b_j,1)$=frontier cost function, with price of input 1 arbitrarily set to unity.

where:

TC_s^a = total actual costs;
TC_s^m = minimum cost possible, given the frontier cost function and actual prices,
$= C(p_{si}^a, y_s)$;
TC_s^b = total costs reflecting technical efficiency but allocative inefficiency; and
λ_s^P = total productive efficiency.

Because there are difficulties sometimes in obtaining convergence of the system of nonlinear equations with the Kopp and Diewert (1982) approach, we have applied as a supplementary step in the results section a linear approximation for TC_s^b suggested by Mensah (1992b). Essentially, it can be approximated as:

$$TC_s^b \cong C(p_{si}^*, y_{sk}), \tag{3}$$

where:

$$p_{si}^* = \frac{CS_{si}^m}{CS_{si}^a} * p_{si}^a, \text{ and} \tag{4a}$$

$$CS_{si}^m = \gamma_i + \sum_{j}^{n} \alpha_{ij} \ell n(p_{sj}) + \sum_{k}^{m} \beta_{ik} \ell n(y_{sk}), \tag{4b}$$

where:

p_{si}^* = optimal cost share of input i based on actual prices; and
p_{si}^a = actual price of input i for local unit s.

The economies of scale for a given local unit s with respect to output k may be defined as:

$$SE_{sk} = 1 - \delta \ell n C(y_{sk}, p_{si}) / \delta \ell n y_{sk} \tag{5a}$$

$$= 1 - \left[\phi_k + \sum_{i=1}^{n} \beta_{ik}(p_{si}) + \sum_{k=1}^{m} \omega_{jk}(y_{sj}) \right], \tag{5b}$$

where all parameters are defined as in equation (1a).

A natural interpretation of SE_{sk} in percentage terms is the degree of scale inefficiency, with values above (below) unity signifying units operating at decreasing (increasing) returns to scale range, and unity values indicating constant returns to scale (Christensen and Greene 1976).

To derive some idea of the degree of institutional X-inefficiency, the cross-sectional average of the input prices must first be derived. The index of input price advantage (IPA) can then be derived as:

$$IPA_s = TC_s^m / TC_s^{m'}, \tag{6}$$

where:

$TC_s^{m'} = C(\bar{p}_i, y_{sk})$,
= frontier function evaluated at cross-sectional average input prices $(\bar{p}_i)$.

The *IPA* as expressed in equation (6) is above (below) unity if a unit's current prices are above (below) the cross-sectional average. The departure from unity indicates the magnitude of possible savings (cost increases) given a movement toward unity. Consequently, the gap between the *IPA* for the *s*th unit and the lowest *IPA* value calculated measures the potential savings to the system (as a percentage of TC_s^m) if the average price level of unit *s* could be reduced that far down.

The DEA Budget Model

The DEA model is a nonparametric technique originally developed by Charnes et al. (1978) to utilize only input and output quantities in deriving a measure of technical efficiency. The technique has subsequently been extended and refined by Banker (1984), Banker et al. (1984), and Banker and Maindiratta (1986) among others. Although the DEA model does not need input price data to derive technical efficiency scores for individual units, such data are necessary for allocative efficiency scores.

To facilitate a comparison of the DEA and the translog results, we modified slightly the DEA approach used by Ferrier and Lovell (1990).[7] Essentially, the dual form of the DEA model presented in Banker et al. (1984) was modified to solve directly for TC_s^m and TC_s^b for each unit. To solve for TC_s^m, the linear program to be solved can be written as:

$$\text{Minimize} \quad TC_s^m = p_{s1}^a \hat{x}_{s1} + p_{s2}^a \hat{x}_{s2} + \ldots + p_{sn}^a \hat{x}_{sn}, \tag{7}$$

subject to:

$$\text{(a)} \quad \sum_{k=1}^{S} \mu_{kj} y_{kj}^a \geq y_{sj}^a \qquad j=1,\ldots,m \text{ outputs;}$$

$$\text{(b)} \quad \sum_{k=1}^{S} \mu_{kh}(-x_{kh}^a) + \hat{x}_{sh} \geq 0 \qquad h=1,\ldots,n \text{ inputs;}$$

$$\text{(c)} \quad \sum_{k=1}^{S} \mu_k = 1; \text{ and}$$

$$\text{(d)} \quad \mu_k \geq 0;$$

where:

S = the number of observations in the sample;
y_{sj}^a = actual output quantities generated by observation s;
x_{sh}^a = actual input quantities used by observation s; and
$\hat{x}_{sh}$ = the optimal input quantities to be determined.

To derive the total costs at the technically efficient, but allocatively inefficient, point (TC_s^b), the following set of $n-1$ restrictions are added to the linear program in equation (7):

---

[7] The first DEA allocative efficiency model was developed by Banker (1980) and extended in Banker and Maindiratta (1988). A different version has been adopted in this study for pedagogical reasons to show the direct link between the translog and DEA cost minimization approaches. See Banker (1980) for an alternative and perhaps simpler approach to the DEA model presented in equation (8).

$$\text{Minimize} \quad TC_s^b = p_{s1}^a \hat{x}_{s1} + p_{s2}^a \hat{x}_{s2} + \ldots + p_{sn}^a \hat{x}_{sn} \tag{8}$$

subject to the same constraints (a) through (d) as for equation (7), as well as:

$$\begin{array}{llll}
\text{(e)} & \vartheta_1 \hat{x}_{s1} & -\hat{x}_{s2} \ldots\ldots\ldots & -\hat{x}_{sn} = 0, \\
\text{(f)} & -\hat{x}_{s1} & \vartheta_2 \hat{x}_{s2} \ldots\ldots\ldots & -\hat{x}_{sn} = 0, \text{ and} \\
\text{(g)} & -\hat{x}_{s1} & -\hat{x}_{s2} \ldots \vartheta_{n-1} \hat{x}_{sn-1} & -\hat{x}_{sn} = 0,
\end{array}$$

where:

$$\vartheta_i = \left(\sum_{k=1}^{S} x_{ki}^a / x_{si}^a \right) - 1 \qquad i = 1, \ldots, n-1 \text{ inputs.}$$

The linear program in equation (8) is equivalent to the usual dual DEA form except that, instead of solving for the technical efficiency factor (λ_s^T), the program solves for the total cost that corresponds to the given level of technical efficiency. The difference between actual total cost (TC_s^a) and the total cost derived from the linear program in equation (8) is the cost difference from technical inefficiency (in percentage terms: $\lambda_s^T = TC_s^b / TC_s^a$). The effect of the additional conditions imposed in equation (8) is to ensure that the input mix at the technically efficient point is the same as the actual mix of inputs used. Thus, the difference in cost of models (7) and (8) measures the result of the improper selection of inputs, given the actual input price structure (i.e., allocative inefficiency; in percentage terms, $\lambda_s^A = TC_s^m / TC_s^b$).

Since different results can be obtained depending on how the production frontier is defined, we estimated two forms of the DEA model: a constant returns-to-scale (CRS) version, and a variable returns-to-scale (VRS) version. The models in equations (7) and (8) are the VRS version; the CRS version can be derived by merely deleting constraint (c) from both equations. Since the VRS version provides the closest fit to the observations, the degree of efficiency identified is usually higher (Ferrier and Lovell 1990). The difference between the two versions is attributable to the fact that the VRS version suppresses scale efficiency from the computed technical and allocative efficiencies; in contrast, the CRS version does not make this distinction. Thus, it is possible to identify the degree of scale inefficiency by comparing the technical efficiency measures derived from the VRS and CRS versions. This approach is preferable if the objective is to compute the effect of scale efficiency in dollar terms.[8]

However, to facilitate comparison with the measure of economies of scale from the translog model, an alternative way of evaluating scale efficiency can be used: compute the sum of the μ_s from the DEA (CRS) results in equation (8). Specifically, if $\Sigma\mu_s > 1$ (< 1), then decreasing (increasing) economies of scale exist for that observation (Banker 1984; Banker and Thrall 1992). Moreover, the magnitude of the departure from unity conveys information about the magnitude of the perceived scale efficiency, with higher or lower values denoting greater inefficiency in one direction or the other.[9]

---

[8] Common usage in the literature refers to the CRS version of the DEA as the "CCR model" and the VRS version as the "BCC model" after the respective authors. The terms CRS and VRS have been adopted because they are more descriptive of the basic difference between the two models.

[9] See Banker et al. (1984, 1090) for an alternative, normalized measure of the returns to scale in the DEA context.

Table 1
Results of Regressions to Derive Translog Cost Functions

	Full Translog		*Parsimonious Model*		*Full Translog Minus 4 Largest Units (Jail #1–4 Deleted)*	
Parameters	*Coefficient*	*T-ratio*	*Coefficient*	*T-ratio*	*Coefficient*	*T-ratio*
Intercept (original)	1.625	2.816	0.593	8.211	1.932	26.013
Intercept (restated)						
LP	−0.003	−0.075	0.509		−0.034	−1.893
BP	0.753	28.829	0.785	51.835	0.383	23.024
MP	0.250	4.384	0.215	14.202	0.651	102.508
JD	0.908	8.451	0.991	238.531	1.352	74.927
AT	−1.386	−8.290	−0.881	−15.019	−2.332	−102.978
TE	−0.148	−0.989			−1.618	−77.007
LPLP	0.218	40.343	0.212	106.122	0.339	114.328
LPBP	−0.186	−46.035	−0.187	−51.692	−0.232	−80.966
LPMP	−0.032	−6.016	−0.025	−8.526	−0.107	−95.493
BPBP	0.206	57.201	0.205	75.070	0.186	56.225
BPMP	−0.020	−6.530	−0.018	−7.191	0.045	37.261
MPMP	0.052	6.754	0.043	9.046	0.062	88.429
LPAT	0.209	22.233	0.200	57.001	0.368	109.234
LPTE	−0.197	−24.148	−0.197	−66.622	−0.128	−48.569
BPAT	−0.198	−31.286	−0.200	−57.001	−0.204	−52.991
BPTE	0.193	38.304	0.197	66.622	0.188	65.039
MPAT	−0.011	−0.884			−0.164	−72.176
MPTE	0.004	0.330			−0.059	−50.636
ATAT	0.268	7.688	0.238	7.317	0.226	57.746
ATTE	−0.131	−3.601	−0.226	−12.653	−0.331	−76.581
TETE	0.206	4.090	0.152	6.776	−0.177	−35.328
LPJD	−0.004	−1.190			−0.088	−45.155
BPJD	0.003	1.197			0.070	32.652
MPJD	0.002	0.332			0.018	25.544
JDAT	0.027	1.253			0.102	38.643
JDTE	−0.009	−0.400			0.225	73.812
JDJD	0.005	0.588			−0.072	−26.666
Weighted system R^2	0.998		0.998		0.999	
Degrees of Freedom	78		87		66	

Note: LP=labor unit price; BP=boarding unit price; MP=maintenance unit price; JD=jail days (output measure 1); AT=average term (output measure 2); TE=technology (proxy measure of capital-labor ratio). Interaction terms are represented by compound expressions such as LPBP, etc.

Finally, just as for the translog model, the degree of institutional X-inefficiency can be assessed by substituting the cross-sectional average price of the inputs in equation (7) to derive $TC_s^{m'}$. The *IPA* for any unit can then be constructed as the ratio $TC_s^{m}/TC_s^{m'}$.

III. Comparison of Translog and DEA Budget Models

Decomposition of Productive Efficiency

The results obtained when the translog cost function was estimated are presented in the first two columns of table 1. Not all the terms in the full translog model were statistically significant, and a formal F-test could not reject a null hypothesis that those insignificant terms were jointly zero at a probability level of below 0.01. Thus, a parsimonious translog model was estimated by dropping all variables found statistically insignificant in the previous analysis. Finally, we estimated the model again with the four largest units deleted. Those results are also presented in table 1.

The correlation to the OLS intercept as a result of shifting the intercept until all the residuals except one were positive is given in table 1, and the resulting total efficiency scores in table 2. For the translog, the unique support point (or pivot) was found to be unit 29 which therefore ensures a 100 percent overall efficiency rating for that unit. The decomposition of the overall productive efficiency rating for all the units in the sample when using the Kopp and Diewert (1982) method is given in the first three columns of table 2.[10]

Table 2 also presents the results obtained when the CRS and VRS versions of the DEA model were used to generate efficiency scores for the individual units. Note that the productive efficiency scores for the same unit are different with the CRS and VRS models, often by substantial amounts. They coincide, however, for specific units, namely, units 10, 14, 16 to 18, 25, and 29, all of which have perfect technical and allocative efficiency scores under the CRS version.

Regardless of the technique used, the degree of allocative inefficiency is lower than the degree of technical inefficiency in this particular sample. When using the translog model, the exact (approximate) solution shows an average allocative efficiency rating of 99 (97.6) percent versus technical efficiency of 93.6 (94) percent. For the CRS (VRS) version of the DEA model, the average ratings are 88.5 (94.1) versus 81.2 (90.7) for allocative and technical efficiency, respectively. In other words, even if the local managers appear to have adopted cost-minimization behavior, as concluded by Hayes and Millar (1990), there is room for improvement for specific local managers. Although these results may not be generalizable beyond this particular sample, the decomposition of the overall productive efficiency as part of a post-implementation exercise would generally provide a central administrator with useful information on where cost savings can be realized in future budgets.[11]

Since it is apparent from table 2 that the DEA and translog models present very different views of the performance of individual local managers, it is necessary to examine the differences in some detail. Table 3 presents an analysis of the product-moment and rank-order correlations between the different sets of efficiency measures.

For the translog model, the exact and approximate results are compared by using those observations for which a solution could be derived with the exact approach. As

---

[10] One of the difficulties encountered in using the exact approach is that sometimes the system of nonlinear equations to be solved does not converge. This may occur if the price of one of the inputs at the optimum is required to be negative (which is inadmissible, given the natural log expression) or if the frontier function is not quasi-concave over some range. For the data used in this study, the exact solution could not be obtained in nine of the 33 cases. The approximate solution was obtained in all cases.

[11] It is important to note that the data used in this study do not include capital costs. Thus, in both this study and the original Hayes and Millar (1990) study, conclusions about possible cost savings from elimination of allocative inefficiency must be qualified. The reason is cogently given by the following observation by Forsund et al. (1980, 22) in a survey of frontier cost and production functions: "suppose . . . the cost-minimizing strategy of the firm dictates that a new plant be built only occasionally, and . . . once a plant is built, certain input substitutions are impossible until the next plant is built. If certain input prices change, such a firm will be (some might prefer to say 'appear to be') allocatively inefficient during any particular year of operation. But to say this raises cost is wrong; such a statement ignores the costs of adjustment which made the firm's strategy optimal."

Table 2
Decomposition of Productive Efficiency into Technical and Allocative Components

	Parsimonious Translog Model					DEA Models						
		Exact Solution		Approximate Solution		CRS			VRS			
Jail No.	*Total Productive Efficiency*	*Technical Efficiency*	*Allocative Efficiency*	*Technical Efficiency*	*Allocative Efficiency*	*Technical Efficiency*	*Allocative Efficiency*	*Productive Efficiency*	*Technical Efficiency*	*Allocative Efficiency*	*Productive Efficiency*	*Capacity Utilization*
1	0.951	0.947	1.004	0.954	0.997	0.600	0.709	0.425	1.000	1.000	1.000	0.873
2	0.894	0.893	1.001	0.895	0.999	0.606	0.669	0.405	1.000	0.944	0.944	0.950
3	0.917	N/C	N/C	0.969	0.947	0.984	0.775	0.763	1.000	1.000	1.000	0.908
4	0.944	0.991	0.952	0.977	0.966	0.762	0.581	0.443	1.000	0.620	0.620	0.865
5	0.804	N/C	N/C	0.885	0.909	1.000	0.844	0.844	1.000	1.000	1.000	0.696
6	0.909	0.913	0.996	0.912	0.997	0.892	0.871	0.777	1.000	1.000	1.000	0.952
7	0.934	0.934	1.000	0.935	0.999	1.000	0.970	0.970	1.000	1.000	1.000	0.850
8	0.913	0.912	1.001	0.913	1.000	0.757	0.779	0.589	0.768	0.860	0.660	0.868
9	0.983	0.983	1.000	0.983	1.000	0.939	0.725	0.681	1.000	0.750	0.750	0.802
10	0.912	0.913	1.000	0.913	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.689
11	0.906	N/C	N/C	0.979	0.925	0.687	0.872	0.599	0.746	0.899	0.671	0.359
12	0.904	0.913	0.991	0.912	0.991	0.856	0.951	0.814	0.860	0.969	0.833	0.725
13	0.926	0.929	0.996	0.928	0.997	0.923	0.907	0.838	0.928	0.932	0.865	0.863
14	0.950	0.950	1.000	0.951	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.840
15	0.892	0.909	0.981	0.905	0.985	0.405	0.946	0.383	0.496	0.953	0.473	0.426
16	0.887	0.942	0.941	0.937	0.947	0.586	0.867	0.507	0.599	0.937	0.561	0.515
17	0.986	0.987	0.999	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	0.533
18	0.906	0.917	0.987	0.913	0.992	1.000	1.000	1.000	1.000	1.000	1.000	0.729
19	0.865	N/C	N/C	0.919	0.942	1.000	1.000	1.000	1.000	1.000	1.000	0.780
20	0.839	N/C	N/C	0.898	0.935	0.702	0.908	0.637	0.744	0.903	0.672	0.828
21	0.903	N/C	N/C	0.911	0.990	0.959	0.947	0.908	0.979	0.974	0.954	0.422
22	0.930	0.930	1.000	0.930	1.000	0.478	0.991	0.474	0.646	0.961	0.621	0.643
23	0.940	0.941	0.999	0.941	1.000	0.540	0.878	0.474	0.583	0.988	0.576	0.609
24	0.928	0.928	1.000	0.928	1.000	0.805	0.971	0.782	0.920	0.937	0.863	0.835
25	0.922	N/C	N/C	0.927	0.995	1.000	1.000	1.000	1.000	1.000	1.000	0.427
26	0.901	0.909	0.991	0.908	0.992	0.531	0.972	0.516	0.715	0.995	0.712	0.630
27	0.935	N/C	N/C	0.938	0.996	0.809	0.810	0.656	1.000	0.789	0.789	0.817
28	0.976	0.972	1.005	0.977	1.000	0.975	0.931	0.907	0.981	0.926	0.909	0.640
29	1.000	1.009	0.991	1.009	0.991	1.000	1.000	1.000	1.000	1.000	1.000	0.548
30	0.894	N/C	N/C	1.162	0.769	0.820	0.956	0.783	0.962	0.924	0.889	0.737
31	0.901	0.916	0.984	0.910	0.991	0.916	0.960	0.880	1.000	1.000	1.000	0.793
32	0.928	0.995	0.932	0.971	0.955	0.668	0.684	0.457	1.000	0.778	0.778	0.845
33	0.836	0.838	0.998	0.839	0.996	0.584	0.739	0.432	1.000	1.000	1.000	0.627
Mean (subset)	0.926	0.936	0.990	0.934	0.991	0.784	0.879	0.698	0.896	0.940	0.840	0.735
Mean (overall)	0.916			0.940	0.976	0.812	0.885	0.726	0.907	0.941	0.853	0.716

Note: N/C = nonconvergence of system of nonlinear equations. The mean subset refers to units for which the exact solution could be derived for the translog model. The overall mean refers to all observations. CRS = constant returns to scale; VRS = variable returns to scale.

Table 3
Correlation Coefficients Between Translog and DEA Measures of Efficiency

	Number of Units	*Pearson Product-Moment Correlation*	*Spearman Rank-Order Correlation*
Translog Model: Exact vs. Approximate Results:			
Technical Efficiency	24	0.988***	0.964***
Allocative Efficiency	24	0.982***	0.820***
DEA Model: CRS vs. VRS Results:			
Technical Efficiency	33	0.731***	0.622***
Allocative Efficiency	33	0.646***	0.507***
Total Productive Efficiency	33	0.684***	0.645***
Translog (Exact) vs. DEA (CRS) Results:			
Technical Efficiency	24	0.343*	0.382*
Allocative Efficiency	24	0.349*	0.016
Total Productive Efficiency	24	0.410**	0.367*
Translog (Exact) vs. DEA (VRS) Results:			
Technical Efficiency	24	0.198	0.271
Allocative Efficiency	24	0.523***	0.055
Total Productive Efficiency	24	0.130	0.172
Translog (Approximate) vs. DEA (CRS) Results:			
Technical Efficiency	33	0.166	0.184
Allocative Efficiency	33	0.013	0.131
Total Productive Efficiency	33	0.183	0.171
Translog (Approximate) vs. DEA (VRS) Results:			
Technical Efficiency	33	0.136	0.123
Allocative Efficiency	33	0.099	0.064
Total Productive Efficiency	33	0.027	0.045
Capacity-Utilization Rate vs. Alternative Technical Efficiency Measures:			
DEA (CRS) Technical Efficiency	33	0.144	0.038
DEA (VRS) Technical Efficiency	33	0.420**	0.385**
Translog (Approx.) Technical Efficiency	33	−0.025	0.004
Translog (Exact) Technical Efficiency	24	−0.049	−0.043

* Significant at a probability level of below 0.10 (one-tailed test).
** Significant at a probability level of below 0.05 (one-tailed test).
*** Significant at a probability level of below 0.01 (one-tailed test).

the results show, the correlation between these observations is very high, with the product-moment and rank-order correlation for allocative efficiency as high as 0.988 and 0.964, respectively; for technical efficiency, the respective correlations are 0.982 and 0.820. In conclusion, the approximate approach appears to be a reasonable proxy for the exact approach without requiring time-consuming calculations.

The comparison of the two versions of the DEA model shows that the productive, technical, and allocative efficiencies are highly and positively correlated. These correlations range around 0.6 and appear to preserve the rank order of the observations to some degree. As discussed earlier, the observed differences in the magnitudes of the technical efficiencies reflect the influence of scale inefficiency. Consequently, preliminary indications are that the degree of scale inefficiency is quite high particularly for units 1 through 4. On a short-term basis, it does not seem reasonable to hold local managers responsible for that part of the overall inefficiency attributable to economies of scale. This suggests that the VRS model should be used for routine post-implementation evaluation of performance so that the focus will be on the "pure" technical and allocative efficiency ratings. The CRS version is still necessary when information is needed on effects of scale efficiency.

A comparison of the consistency of the DEA and translog models shows that the translog model is generally in greater agreement with the CRS version of the DEA model than with the VRS version. This result is consistent with the observation that the VRS model provides a close fit to the data, while both the translog and CRS models do not envelop the data as closely. This is evident from table 2 where 19 of the 33 units were rated by the VRS model as operating on the production frontier (i.e., 100 percent technical efficiency rating). In contrast, the CRS version picked ten units as technically perfect while the exact translog solution identified only one.[12]

These results are consistent with those from two previous empirical studies that compared the translog to the DEA model. Banker et al. (1986), in a study of hospitals in North Carolina, found a weakly significant level of positive correlation between productive efficiency computed from the translog and technical efficiency computed from a DEA model. In their study of bank operations, Ferrier and Lovell (1990) found a positive but statistically insignificant correlation between the technical efficiencies computed from the translog and various DEA models. Both studies concluded that the DEA model provided a closer fit to the data than the translog and that this factor accounted for the difference in results.

Validation of Technical Efficiency

Although the comparisons up to this point are interesting in their own right, they do not address the issue of which set of efficiency measures should be relied on as the best guide to action. To provide some insight into this area, the observed technical efficiencies calculated for each method and model can be compared to the degree of capacity utilization in conformance with reasoning advanced by Banker et al. (1986). Essentially, the argument for this validation approach rests on the notion that, unless the operating costs are truly variable over all ranges, the degree of capacity utilization should affect the perceived technical efficiency of the individual units. That is, to the extent that the fixed resources are under-utilized in generating output, an inefficiency exists and is of the kind that a sensitive technical efficiency measure should reflect. Note that additional inefficiencies unrelated to capacity utilization may exist; thus the correlation between capacity utilization and technical efficiency is not expected to be perfect, but only to exist. The results of this validation attempt are shown at the bottom of table 3.

As is evident, only the DEA (VRS) model's technical efficiency scores show a statis-

---

[12] The rather anomalous result of having units with efficiency ratings above 100 percent stems from the choice of the pivot or support point. Changing the pivot point by shifting the intercept of the deterministic frontier will rescale the technical efficiency measures with no effect on the relative rankings.

Table 4
Estimated Degrees of Scale Inefficiency

	Measures of Size		*Translog*		
Units	*Bed Capacity*	*Jail Days*	*Full Model*	*Parsimonious*	*DEA (CRS) Results*
All data used					
Units 1–4	633	207,799	1.005 DRS	0.991 CRS**	17.344 DRS
Units 5–10	90	27,623	0.983 IRS	0.991 CRS**	1.661 DRS
Units 11–16	70	15,948	0.983 IRS	0.991 CRS**	0.844 IRS
Units 17–22	55	13,271	0.984 IRS	0.991 CRS**	0.858 IRS
Units 23–28	35	8,219	0.970 IRS	0.991 CRS**	0.806 IRS
Units 29–33	14	3,666	0.982 IRS	0.991 CRS**	0.677 IRS
Largest units (1–4) deleted					
Units 5–10	90	27,623	0.978 IRS		1.661 DRS
Units 11–16	70	15,948	1.004 CRS/DRS		0.856 IRS
Units 17–22	55	13,271	1.073 DRS		0.935 IRS
Units 23–28	35	8,219	0.963 IRS		0.693 IRS
Units 29–33	14	3,666	0.950 IRS		0.798 IRS

Notes: Results for the DEA model are multiples of a CRS unit; thus values less than (above) unity imply a current size smaller (larger) than optimal for perfect scale efficiency. An identical interpretation applies to the translog results. For the translog model with all data used, an F-ratio shows that a null hypothesis of constant returns to scale cannot be rejected at any reasonable probability level. CRS = constant results to scale; IRS = increasing returns to scale; DRS = decreasing returns to scale; and jail days = output measure 1.

tically significant relationship with the degree of capacity utilization. The DEA (CRS) model's scores show a positive but statistically insignificant relationship, while the results for the translog model have the wrong sign in three of the four comparisons, although all of them are statistically insignificant.

These results confirm the previous conclusion that the VRS model should perhaps be adopted exclusively for routine performance evaluation in a budgeting context. Although the use of this version may result in more local units being identified as efficient, those units identified as inefficient can easily be proved to be so relative to their closely matched counterparts. Consequently, a local manager will have virtually no basis to dispute a finding of inefficiency. In a budgetary context, it is important for subordinate managers to perceive their performance as having been evaluated fairly.

Comparison of Scale Efficiencies

As mentioned earlier, the analysis of possible differences in economies of scale in different segments of the observations can provide information useful for long-range budgeting decisions. Since both the translog and DEA models can generate such information, it is worthwhile to compare the results obtained from the two models. These are presented in table 4.

Results are presented for the full translog model, the parsimonious translog model, and the DEA (CRS) model. Although two output measures (plus a technology indicator) were included in the translog and DEA models, the results in table 4 are reported only for jail days, the only output measure that conveys a volumetric effect and is related to jail (bed) capacity.

If the statistical insignificance of some of the coefficients in the full translog is ignored, the estimated scale inefficiencies computed for units in the various ranges identified show that the data cover a wide spectrum, with both decreasing and increasing returns to scale. For the full translog model, units 1 through 4 (the largest in terms of bed capacity; see col. 2 of the appendix) were operating (on average) in a range marked by some minor degree of decreasing returns to scale, while the other units were operating in a range subject to increasing returns to scale. When the statistical significance of the coefficients in the full translog model is considered, however, this conclusion no longer holds. The parsimonious translog results show that the perceived differences in economies of scale are statistically insignificant. Thus, the translog function does not permit rejection of an assumption of constant returns to scale across all segments.

In contrast, the DEA model suggests much greater differences in the degree of economies of scale prevailing in the different segments. Results for units 1 through 4 and units 5 through 10 show evidence of scale inefficiency, and the computed coefficients of 17.344 and 1.661, by group, suggest that the units are, on average, 17 to 1.6 times larger than the optimal size range. In contrast, units 11 through 16 average 84.4 percent of the optimal size, while the smallest units (29–33) average 67.7 percent of the optimal size.[13] The statistical significance of these scale effects can be evaluated formally with the translog approach, but the development of asymptotic tests has unfortunately not progressed sufficiently at present for such an evaluation for the DEA model, although work is currently underway to redress this deficiency (see e.g., Banker 1992). However, the magnitudes of the differences suggest that they are material in an accounting sense.

Up to this point, the results are similar to those reported by Banker et al. (1986), who found the translog to be insufficiently sensitive in detecting differences in economies of scale. In contrast, Ferrier and Lovell (1990) found the translog and the DEA techniques to be consistent in their evaluation of the degree of economies of scale. To gain further insight, we decided to evaluate the issue further. Since units 1 through 4 in this sample are obviously much larger than the other units, they were treated as outliers, and the data were reanalyzed to determine the subsequent effect on the results.

The results of this reanalysis on the translog function parameters are presented in the last set of columns in table 1. Note that, unlike the results when all observations were used, all coefficients for the full translog are now statistically significant. Thus. a null hypothesis of a constant returns to scale (with respect to jail days) can be rejected at any reasonable probability level.

The average economies of scale with units 1 through 4 deleted are given in table 4. The translog model now shows units 5 through 10 as experiencing increasing returns to scale, while units 11 through 15 and 17 through 22 are experiencing either constant or decreasing returns to scale. Units 23 through 28 and 29 through 33 continue to experience increasing returns to scale. In contrast, the results with the DEA model are virtually unchanged. In our view, the relative stability of the DEA results makes it preferable for inferring the degree of scale inefficiency.

---

[13] The optimal size ranges differ because neither the second output measure (AT) nor technology (TE) can be controlled for simultaneously in computing the effect of economies of scale with the DEA model.

Institutional X-Efficiency

As discussed previously, X-efficiency theory suggests that the largest potential savings may lie in questioning the basic institutional assumption that the local units are price-takers unable to influence their input unit prices. To the extent that a central administrator can pursue policies that influence input prices (such as centralized purchasing policies, etc.), it is worthwhile to assess the effect of this assumption of exogeneity of input prices on the observed results. Table 5 presents the results obtained when *IPA* is computed with the translog and DEA (VRS) models.

As the results show, in contrast to the previous conflict in results, the *IPA*s computed under the translog and DEA (VRS) models are in very close agreement. They have identical means of 0.91, and the correlation coefficient between them is 0.989. Both also show the same wide disparity in *IPA*s, which range from 0.24 (unit 27) to 1.93 (unit 3) with the translog model, and from 0.37 to 1.80 for the DEA (VRS) model. These results suggest that a sustained program designed to change the institutional framework that supports this huge cost differential might pay off in reducing costs for the entire system.

IV. Evaluation of Comparative Merits

Effect of Outliers on Productive Efficiency

The conflicting results from the DEA and translog budget models in the measurement of productive efficiency suggest a need to examine possible causes. One obvious culprit suggested in the literature is the effect of outliers. In some previous reviews of the literature (e.g., Bauer 1990; Forsund et al. 1980), the DEA model was presumed to be more sensitive to the effect of outliers than the translog model. The issue has not been specifically addressed, but we are skeptical about this presumption for the following reason. Since the translog model develops parameter values that apply to all observations, outliers are more likely to affect the resulting parameter values, and consequently the respective efficiency scores derived. In addition, for full frontier translog functions, if the point chosen as the pivot happens to be an outlier, the resulting technical efficiency scores for all observations will be affected. In contrast, the effect of outliers on the efficiency scores in DEA is likely to be localized to the immediate observations for which the outlier forms a basis. Outliers that lie to the interior of the production possibility set will have no effect whatsoever on the efficiency scores of other observations.

To evaluate the effect of outliers on the estimated productive efficiencies, unit 29, which was the support point for the translog and a key support point for the DEA results, was defined as an outlier. The translog model was then reestimated, and the resulting frontier function was used to calculate new technical and allocative efficiency measures. A similar procedure was carried out with the DEA (VRS) model. The rank-order and product-moment correlations between the original set of efficiency measures and the revised measures are presented in table 6.

The results show that the removal of unit 29 has a more profound effect on the efficiences from the translog model than on the DEA model. Although there is no significant shift in the means of the overall productive, technical, and allocative efficiency scores, the correlations between them tell a different story. For both the product-moment and rank-order correlation coefficients, the deletion of unit 29 has virtually no effect on the DEA results, with all correlations above 0.98. In contrast, the deletion seriously affects the relative rankings of units under the translog approach; even the

Table 5
Estimates of Potential Savings From Achievement of Institutional X-Efficiency Through Input Price Reductions

	Translog Model			*DEA (VRS) Model*		
Jail No.	*Minimum Cost at Current Input Prices*	*Minimum Cost at Average Input Prices*	*Index of Input Price Advantage*	*Minimum Cost at Current Input Prices*	*Minimum Cost at Average Input Prices*	*Index of Input Price Advantage*
1	$12,249,998	$9,302,652	1.32	$12,882,637	$9,987,895	1.29
2	5,662,397	6,135,808	0.92	5,979,933	6,557,884	0.91
3	2,933,661	1,516,910	1.93	3,198,450	1,780,285	1.80
4	2,467,112	1,725,607	1.43	1,620,063	1,148,231	1.41
5	228,815	283,667	0.81	284,631	377,263	0.75
6	477,452	425,087	1.12	525,161	471,230	1.11
7	321,344	354,200	0.91	344,065	380,750	0.90
8	491,261	488,739	1.01	355,272	371,185	0.96
9	349,760	386,460	0.91	266,966	299,228	0.89
10	128,529	280,357	0.46	140,858	316,407	0.45
11	135,544	181,853	0.75	100,367	134,060	0.75
12	250,251	235,601	1.06	230,616	213,956	1.08
13	227,950	281,579	0.81	212,869	255,964	0.83
14	172,183	222,708	0.77	181,151	234,339	0.77
15	194,485	268,569	0.72	103,134	142,586	0.72
16	254,595	255,382	1.00	160,995	161,157	1.00
17	147,525	158,915	0.93	149,586	162,768	0.92
18	253,174	186,153	1.36	279,551	205,413	1.36
19	167,053	172,570	0.97	193,069	225,945	0.85
20	220,550	236,606	0.93	176,591	188,399	0.94
21	100,025	116,186	0.86	105,728	125,440	0.84
22	225,479	234,270	0.96	150,575	136,460	1.10
23	145,236	215,232	0.67	88,980	132,095	0.67
24	81,528	211,840	0.38	75,784	175,034	0.43
25	94,350	95,222	0.99	102,317	105,589	0.97
26	120,446	171,735	0.70	95,185	121,442	0.78
27	75,975	314,905	0.24	64,108	173,232	0.37
28	94,953	117,279	0.81	88,378	108,527	0.81
29	105,993	89,902	1.18	105,993	90,478	1.17
30	92,773	95,803	0.97	92,197	95,375	0.97
31	59,184	97,139	0.61	65,659	108,937	0.60
32	94,802	108,362	0.87	79,489	90,478	0.88
33	62,519	76,655	0.82	74,761	90,478	0.83
Mean	$869,300	$758,908	0.91	$865,913	$762,682	0.91

Notes: The index of input price advantage (*IPA*) measures the increase or decrease in total costs attributable to current input prices higher or lower than the cross-sectional average of the period. Values below (above) unity indicate lower (higher) actual input prices relative to the average, and a consequent input price advantage (disadvantage). The product-moment correlation between the two sets of *IPA*s is 0.989.

overall productive efficiency ratings are changed since the rank-order correlation between the original and revised scores are only 0.612, as compared to 0.998 for the DEA model.[14] Overall, a conclusion that the DEA model results are relatively more impervious to deletion of outliers than the full frontier translog model is justified.

---

[14] A similar comparison by deleting units 1 through 4 could not be performed because the monotonicity condition (Berndt 1991, 469–76) was violated by the translog function estimated; i.e., the fitted cost shares were not positive for all inputs for all observations. In fact, this result suggests that the translog function may not be as widely applicable as the DEA, which does not impose a specific functional form on the data.

Table 6

Correlation of Revised Efficiency Measures with Original Measures

(with removal of outlier unit 29)

	Mean of Measures (Revised)	*Mean of Measures (Original)*	*Product-Moment Correlation*	*Rank-Order Correlation*
Technical Efficiency Measures				
Translog (approximate)	99.09%	93.77%	0.06	0.491
DEA (VRS)	90.64	90.40	0.999	0.981
Allocative Efficiency Measures				
Translog (approximate)	94.95	97.53	−0.522	0.221
DEA (VRS)	94.10	93.88	0.99	0.981
Productive Efficiency Measures				
Translog (approximate)	92.73	91.30	0.343	0.612
DEA (VRS)	85.28	84.81	0.997	0.998

Relative Merits of the Translog Approach

From the analysis so far, it is evident that the DEA is perhaps to be preferred as a budget model. Apart from its greater simplicity, in the one validation test that could be performed, the DEA (VRS) model generated technical efficiency scores that were more highly correlated with the degree of capacity utilization than the translog model. In addition, the DEA (CRS) model generated economies of scale measures that seemed better adapted to the data than the translog model. Moreover, the DEA model is less sensitive to outliers (except in the local region of the outlier), and appears to be more widely applicable than the translog function.[15] Finally, as noted by prior researchers as well as in this study, the DEA provides a closer fit to the data than the translog model. Consequently, budgetary cuts based on estimated savings from the elimination of inefficiencies are less likely to be disputed by subordinate managers than those based on the translog function.

Given the substantial merits of the DEA approach, it is debatable if the translog approach is necessary in this context. The primary argument for using a translog cost function rests on the benefits of standard errors obtainable from the econometric modeling exercise. In not-for-profit settings, the definition and measurement of inputs and outputs often present problems. As an illustration, consider the county jail data used in this study. One could argue that there may be differences in the quality of output as measured, for example, by the rate of recidivism of inmates released from the jails. If such a qualitative factor is theorized to have systematic cost implications, the issue should be examined statistically and either supported or rejected. The translog approach can be used to test whether these are important cost drivers. Once the most suitable definitions of outputs and inputs have been identified with the translog, the DEA approach can then be used to perform all of the routine analyses suggested.

[15] In fairness to the translog approach, it should be noted that it has been applied to a wide variety of situations such as electric power generation (Greene 1980b), railroad operations (Kumbkahar 1988), and hospitals (Banker et al. 1986).

Although Banker (1992) has extended the DEA approach to incorporate formal statistical tests, much work remains to be done in this area. Among other things, further research will be needed to determine the robustness of the statistical tests with the DEA approach compared to the translog. In addition, this study examined the relative performance of only the translog function compared to the DEA approach; it is conceivable that other flexible functional forms such as the generalized Leontieff might have fared better than the translog in this comparison.

V. Conclusions

This article has demonstrated ways in which the translog budget idea initially presented by Hayes and Millar (1990) can be extended. In the process, we have pointed out some problems with Hayes and Millar's criticism of LIB as a defective control device in not-for-profit settings. We have shown that the detailed data on inputs, outputs, and unit costs typically required under LIB are necessary for any effective pre- or post-implementation budgetary control system. The combination of these data with a technique such as the translog or the DEA budget model will permit useful information to be extracted for the short- and long-term objectives of budgeting.

Furthermore, through extensive comparative analyses, we have shown that the DEA budget approach might actually be preferred to the translog approach for routine performance evaluation and other budgeting purposes. The need for the translog model may arise when there are specification issues about how best to represent inputs or outputs. In such cases, the translog model may be used to identify the best proxies. DEA can then be applied to elicit the other items of information illustrated in this study.

Appendix
Raw Data and Summary Details for Tennessee County Jails

			Outputs			*Actual Input Quantities*			*Actual Unit Prices*		
Jail No.	*Bed Capacity*	*Total Cost*	*Jail Days*	*Average Term*	*Technology*	*Employees (Annual)*	*Admissions*	*Total Sq. Ft.*	*Labor Pay (Weekly)*	*Boarding Costs*	*Maintenance Costs*
1	1,244	$12,882,637.35	396,389	5.44662	2.92722	371.0	72,777	71,872	$456.32599	$ 31.20395	$25.15954
2	782	6,336,293.33	271,158	6.59271	2.85730	260.0	41,130	86,583	335.46526	28.41425	7.30080
3	292	3,198,449.96	96,726	14.37663	3.84062	69.0	6,728	39,555	412.45917	238.98986	2.79668
4	212	2,613,817.56	66,922	7.41190	2.32086	79.0	9,029	10,224	420.00896	88.39836	8.82942
5	108	284,631.18	27,429	11.52963	7.51479	10.0	2,379	14,552	296.84996	31.84397	3.74604
6	99	525,161.21	34,390	9.55278	6.28128	15.0	3,600	7,616	458.17472	31.99749	6.90572
7	99	344,065.03	30,732	6.77215	9.35525	9.0	4,538	3,998	443.87613	21.46981	9.73012
8	93	538,034.54	29,461	8.42465	4.24816	19.0	3,497	8,537	375.03423	38.03094	4.04200
9	86	355,821.93	25,185	9.22190	4.60000	15.0	2,731	3,204	298.55297	30.08209	12.73297
10	73	140,858.41	18,347	3.63739	12.56644	4.0	5,044	3,757	376.46947	10.60958	2.40565
11	72	149,636.92	9,437	10.55593	3.23185	8.0	894	7,304	221.18108	55.16758	1.13715
12	72	276,688.27	19,052	7.53043	7.45675	7.0	2,530	4,013	456.49099	29.62562	8.86438
13	71	246,159.68	22,362	6.12658	8.75225	7.0	3,650	3,525	319.76706	27.42326	8.41690
14	70	181,151.03	21,471	9.62825	8.40352	7.0	2,230	2,427	316.49945	22.06806	6.89471
15	68	218,006.09	10,582	4.13037	3.22131	9.0	2,562	4,226	231.42000	29.36931	8.15366
16	68	286,974.14	12,786	5.80654	3.89224	9.0	2,202	2,864	305.52656	29.43852	27.64109
17	66	149,585.67	12,829	8.16614	7.02959	5.0	1,571	1,290	268.37262	40.34826	12.72998
18	64	279,550.97	17,024	7.60679	9.32822	5.0	2,238	2,960	471.50146	54.09295	12.12857
19	58	193,069.24	16,505	16.13392	6.45988	7.0	1,023	7,712	228.56060	83.61645	3.15528
20	53	262,866.29	16,019	7.52066	5.48596	8.0	2,130	4,736	241.71168	45.46121	13.82640
21	48	110,819.98	7,394	7.38661	5.06438	4.0	1,001	2,896	314.96284	36.14260	3.15227
22	42	242,478.03	9,855	3.44340	4.35484	6.2	2,862	2,850	520.26504	17.74436	8.40709
23	42	154,428.96	9,337	5.43481	3.19760	8.0	1,718	2,408	212.34236	30.82676	5.45439
24	37	87,858.20	11,275	3.43960	7.72260	4.0	3,278	2,778	235.32673	9.79344	2.45044
25	37	102,316.61	5,765	11.10790	3.94863	4.0	519	1,908	337.28178	44.84200	4.65881
26	35	133,717.98	8,046	4.25264	4.40877	5.0	1,892	2,960	324.36165	17.20160	5.68869
27	28	81,278.50	8,352	1.37504	7.62740	3.0	6,074	3,914	246.64936	5.83716	1.87692
28	28	97,258.74	6,540	10.48077	3.58356	5.0	624	1,125	215.67004	47.61849	10.19608
29	22	105,993.03	4,398	10.64891	3.01233	4.0	413	636	266.43034	73.95034	31.50005
30	18	103,754.36	4,841	8.73827	3.31575	4.0	554	5,712	312.17269	58.26300	1.14579
31	16	65,659.32	4,631	4.68725	4.22922	3.0	988	1,899	226.29051	23.02530	4.00685
32	10	102,203.49	3,085	5.56859	1.69041	5.0	554	598	273.63685	52.23392	3.54569
33	6	74,761.19	1,374	4.54967	0.94110	4.0	302	1,456	268.59707	40.68212	4.53777

Source: Robert D. Hayes (coauthor in Hayes and Millar 1990).

References

Banker, R. D. 1980. *Studies in Cost Allocation and Efficiency Evaluation.* Ph.D. dissertation, Harvard University, Cambridge, MA.

———. 1984. Estimating most productive scale size using data envelopment analysis. *European Journal of Operational Research* 17 (July): 35–44.

———. 1992. Maximum likelihood, consistency and data envelopment analysis: A statistical foundation. *Management Science* (Forthcoming).

———, A. Charnes, and W. W. Cooper. 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30 (September): 1078–92.

———, R. F. Conrad, and R. P. Strauss. 1986. A comparative application of data envelopment analysis and translog methods: An illustrative study of hospital production. *Management Science* 32 (January): 30–44.

———, and A. Maindiratta. 1986. Piecewise loglinear estimation of efficient production surfaces. *Management Science* 32 (January): 126–35.

———, and ———. 1988. Nonparametric analysis of technical and allocative efficiencies in production. *Econometrica* 56 (November): 35–56.

———, and R. Thrall. 1992. Estimating returns to scale in data envelopment analysis. *European Journal of Operational Research* (Forthcoming).

Barlev, B., and J. L. Callen. 1986. Total factor productivity and cost variances: Survey and analysis. *Journal of Accounting Literature* 5: 35–56.

Bauer, P. A. 1990. Recent developments in the econometric estimation of frontiers. *Journal of Econometrics* 46 (Supplement): 39–56.

Berndt, E. R. 1991. *The Practice of Econometrics: Classic and Contemporary.* New York: Addison Wesley.

Callen, J. L. 1988. An index number theory of accounting cost variances. *Journal of Accounting, Auditing & Finance*, n.s., 3 (2): 89–112.

Caves, D. W., L. R. Christensen, and J. A. Swanson. 1981. Productivity growth, scale economies, and capacity utilization in U.S. railroads, 1955–74. *American Economic Review* 71 (December): 994–1002.

Charnes, A., W. W. Cooper, and E. Rhodes. 1978. Measuring the efficiency of decision-making units. *European Journal of Operations Research* 2 (November): 429–44.

Christensen, L. R., and W. H. Greene. 1976. Economies of scale in U.S. electric power generation. *Journal of Political Economy* 84 (August): 655–76.

Darrough, M. N. 1988. Variance analysis: A unifying cost function approach. *Contemporary Accounting Research* 5 (Fall): 199–221.

Farrell, M. J. 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society, series A*, 120 (3): 253–81.

Ferrier, G. D., and C. A. K. Lovell. 1990. Measuring cost efficiency in banking: Econometric and linear programming evidence. *Journal of Econometrics* 46 (Supplement): 229–45.

Forsund, F. R., C. A. K. Lovell, and P. Schmidt. 1980. A survey of frontier production functions and of their relationship to efficiency measurement. *Journal of Econometrics* 13 (May): 5–25.

Greene, W. H. 1980a. Maximum likelihood estimation of econometric frontier functions. *Journal of Econometrics* 13 (May): 27–56.

———. 1980b. On the estimation of a flexible frontier production model. *Journal of Econometrics* 13 (May): 101–15.

Hayes, R. D., and J. A. Millar. 1990. Measuring production efficiency in a not-for-profit setting. *The Accounting Review* 65 (July): 505–19.

Kopp, R. J. 1981. The measurement of productive efficiency: A reconsideration. *Quarterly Journal of Economics* 96 (August): 477–503.

———, and W. E. Diewert. 1982. The decomposition of frontier cost function deviations into measures of technical and allocative efficiency. *Journal of Econometrics* 19 (August): 319–31.

Kumbkahar, S. C. 1988. On the estimation of technical and allocative inefficiency using stochastic frontier functions: The case of U.S. class 1 railroads. *International Economic Review* 29 (November): 727–43.

Liebenstein, H. 1966. Allocative efficiency versus X-efficiency. *American Economic Review* 56 (June): 397–409.

———. 1976. *Beyond Economic Man: A New Foundation for Microeconomics.* Cambridge, MA: Harvard University Press.

Lynn, E. S., and R. J. Freeman. 1983. *Fund Accounting: Theory and Practice.* Englewood Cliffs, NJ: Prentice-Hall.

Marcinko, D. E., and E. Petri. 1984. Use of the production function in calculations of standard cost variances: An extension. *The Accounting Review* 59 (October): 488–95.

Mensah, Y. M. 1982. A dynamic approach to the evaluation of input-variable cost center performance. *The Accounting Review* 57 (December): 681–700.

———. 1992a. Alternative cost function input characterizations in the evaluation of line-item budgeting. *Journal of Econometrics* (Forthcoming).

———. 1992b. A simplification of the Kopp-Diewert method of decomposing productive efficiency and some implications. Working paper, Rutgers University, New Brunswick, NJ.

Schmidt, P., and C. A. K. Lovell. 1979. Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers. *Journal of Econometrics* 9 (February): 343–66.

Ramanathan, K. V. 1982. *Management Control in Nonprofit Organizations.* New York: John Wiley & Sons.

Zieschang, K. 1983. A note on the decomposition of cost efficiency into technical and allocative components. *Journal of Econometrics* 23 (October): 401–05.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 89–92

A Rejoinder to "Measuring Production Efficiency in a Not-for-Profit Setting: An Extension"

Robert D. Hayes
Tennessee State University
James A. Millar
University of Arkansas

THE "extension" of Hayes and Millar (1990) by Mensah and Li (1993) brings up essentially two areas of controversy. First, Mensah and Li find fault with our "implied" assumption of cross-sectionally fixed cost shares. Second, their use of an additional measure of output (AT) and an alternative measure of technology (TE) raises questions regarding model specification. In this rejoinder, we discuss these two areas of concern in order to address any previous omissions, needed clarifications, or differences in purposes. Because different methods were used in the two investigations, a third issue concerning methodologies seems relevant. Therefore, we also include some general comments regarding "best-practice" and "average-practice" methods of analysis.

I. Cross-Sectional Constancy of Cost Shares

Mensah and Li (1993) apparently find an implied assumption of cross-sectionally fixed cost shares in Hayes and Millar (1990). Although it is seldom possible to be totally successful in communicating complicated concepts, we find the interpretation that "this conclusion relates to the constancy of the optimal cost shares in a cross-sectional context" (Mensah and Li 1993) to be inconsistent with our presentation. How this interpretation follows from Hayes and Millar (1990, 513) is unclear. After all, it is the variation in input prices that allows the estimation of the cost function across operating units. Input prices are given for a particular operating unit, but input prices are not identical across the whole sample.

We believe that it follows from our estimation procedure and hypothesis tests that we specifically reject the homogeneous Cobb-Douglas functional form (prior to testing the feasibility of such a specification) in Hayes and Millar (1990). Moreover, our intended interpretation is that (1) line-item budgeting requires that the cost shares are fixed for the duration of the budget period for a given operating unit, and (2) fixed cost

Invited by the Editor.
Received August 1992.
Revision September 1992.

shares may inhibit the discretionary power of operating unit managers to substitute one input for another in an environment of changing relative input prices. The estimation procedure with cross-sectional data simply contemplates that the budgetary units are operating at different points on an underlying production function (albeit, "average" production function).

II. Measures of Output and Technology

In the corrections environment, operating costs are often quoted on a "per inmate" basis, and we chose the output measure "jail days" accordingly (Hayes and Millar 1990). Mensah and Li (1993) use this same output variable, but elect to use an additional measure of output designated as AT.

In both studies, correctional facility operating characteristics were postulated to affect operating costs through ways other than output alone. For Hayes and Millar (1990) at least, choosing a variable to represent these physical conditions of production was more difficult than choosing an output measure. We found that the variable for percentage of beds occupied is somewhat correlated (0.3917) with the output-level variable (Hayes and Millar 1990, 510). For this reason, we did not use percentage of beds occupied (average daily number of inmates/average daily number of beds available) in our fitted models. Our firsthand knowledge of the data caused us to choose CAP/STAFF (Hayes and Millar 1990, 510) as a proxy for a capital-labor ratio. With this variable as an index of technology, we attempted to capture the effect of differences in technology upon costs. This ratio formulation of a technology index is broad enough to incorporate aspects of capital via number of beds available and aspects of the labor force via number of employees. The exact formulation of "technology" may be questioned and debated, but the need to consider, at least initially, some measure of technology is well established.

Our decision concerning which, and how many, explanatory variables to employ was conditioned by our purpose and by the data. Our purpose included the intent to interpret the results in a line-item budgeting context by using the elasticities derived from the fitted model coefficients (Hayes and Millar 1990, 514–17). Therefore, a primary consideration was the possible effects of multicollinearity on the stability of the estimated coefficients.

Regarding the choice of explanatory variables, Mensah and Li's (1993) use of AT and TE as variables does have the apparent advantage of separating differences in capacity utilization from differences in technology. We, of course, do not argue with the use of any theoretically supportable and statistically manageable measure of output or technology. In fact, efforts to identify alternative measures of output and technology might be fertile. However, for the data set used in both studies, we are reluctant to fully accept Mensah and Li's choice of AT as an additional explanatory variable for "output" because of its 0.62932 correlation with the unit price of boarding (BP). Yet, this correlation is not necessarily a problem since Mensah and Li are not attempting to analyze input-substitution using Allen-Uzawa partial elasticities (Hayes and Millar 1990, 516). We also note that, as expected, Mensah and Li's TE variable is highly correlated (0.86812) with our CAP/STAFF variable. This, of course, presents no problem because, unlike AT and BP, TE and CAP/STAFF are never in the same model. Still, there may be other measures of technology that are as good or better than either TE or CAP/STAFF.

III. "Average-Practice" versus "Best-Practice" Approaches

It might be said that ordinary least squares statistical regressions "average" performance across many observational units and, therefore, explain typical behavior rather than behavior of individual units. This result is partially attributable to incorporating error terms with both positive and negative values, $N(0,\sigma^2)$, in the statistical model. Alternatively, best-practice (actually, relative-best-practice) approaches that follow from the empirical work of Farrell (1957) attempt to estimate "frontier" functions. In best-practice analysis, a routine objective is to determine the minimum cost incurred by a given "firm" to produce a given output (or maximum output for given inputs). An appropriate statistical model for this purpose would incorporate error terms made up of two components: $e_i=v_i+u_i$. One component consists only of error terms with non-negative values: $N^+(0,\sigma^2_{V_i})$. The other component is assumed to be independently and identically distributed: $N(0,\sigma^2_{u_i})$. (See Banker et al. 1987 for a more complete discussion.)

Use of a best-practice approach versus an average-practice approach should be determined according to the purpose of the investigation. Some issues might be more manageable under one approach than under the other. For example, what part does a "free lunch" play during the interpretation phase of analysis in each approach; what additional complications arise in a single-period cross-sectional analysis? The implications of a "free lunch" are clearly different for a best-case scenario versus an average-case scenario. Mensah and Li (1993) point out one aspect of this issue by noting that technical efficiency scores of all observations will be affected if the point chosen as a pivot happens to be an "outlier." Also, interpretation of technical efficiency with respect to the degree of capital utilization seems insufficient when demand conditions are not a part of the analysis (Barlev and Callen 1986, 51).

In Hayes and Millar (1990), we explicitly state that optimal efficiency cannot be demonstrated with our methodology. Specifically, we did not attempt to resolve "... the possibility that a more cost-efficient input combination might have been utilized" (Hayes and Millar 1990, 514). Subsequently, Mensah and Li (1993) adequately demonstrate the utility of the best-practice approach by attempting to resolve questions that we could not address with average-practice methodology.

In summary, the focus of our earlier investigation was on the inadequacy of traditional line-item budgeting, for the typical not-for-profit operating unit, in providing adequate feedback and motivation to encourage cost-minimizing behavior and discourage dysfunctional decision-making by managers. The procedure used in Hayes and Millar (1990) is appropriate for that purpose. However, Mensah and Li (1993) are right on target in pointing out that the translog budget idea initially presented by Hayes and Millar can be extended. Mensah and Li pursue their purpose with a best-practice approach. Their focus is to make statements about the relative performance of individual operating units, about individual managers. The procedure used is appropriate for their purpose. It may also be noteworthy that Mensah and Li relegate some additional interesting points to footnotes.

IV. Conclusion

Hayes and Millar (1990) estimated a short-run cost function because it was unlikely that the institutions in the sample had been able to make optimal adjustments in their physical facilities. Hence, the proper procedure was to estimate the short-run cost function given a reasonable suspicion that the observed units were not operating in long-run equilibrium. Unfortunately, the data set included operating expenditure amounts only. Since Mensah and Li (1993) use this same data set, in our opinion, their "extension" must be considered pedagogical in nature because the data do not encompass capital cost. Other limitations concerning generalizability are appropriate and are given by both Hayes and Millar (1990) and Mensah and Li (1993).

As the segment of society most responsible for reporting outcomes of financial and operational decisions, the accounting profession is responsible to users of that information. That obligation is well served when users can distinguish efficient usages of economic resources from inefficient ones. At present, most not-for-profit systems are designed to ensure dollar (financial) accountability. Much less emphasis is placed on (operational) accountability for physical resources. Do Mensah and Li (1993) offer a more complete budgetary system that potentially enhances, to some degree, users' ability to identify and respond to efficient and inefficient usage of economic resources? We think that, perhaps, they do.

References

Banker, R. D., S. M. Datar, and M. V. Rajan. 1987. Measurement of productivity improvements: An empirical analysis. *Journal of Accounting, Auditing & Finance* 2 (Fall): 319–54.

Barlev, B., and J. L. Callen. 1986. Total factor productivity and cost variances: Survey and analysis. *Journal of Accounting Literature* 5: 35–36.

Farrell, M. J. 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society* (Series A) 120: 253–81.

Hayes, R. D., and J. A. Millar. 1990. Measuring production efficiency in a not-for-profit setting. *The Accounting Review* 65 (July): 505–19.

Mensah, Y. M., and S. Li. 1993. Measuring production efficiency in a not-for-profit setting: An extension. *The Accounting Review* 68 (January): 66–88.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 93–112

On the Optimality of Public Signals in the Presence of Private Information

Michael Alles
University of Texas at Austin
Russell Lundholm
Stanford University

SYNOPSIS AND INTRODUCTION: Security regulation relies heavily on the public disclosure of information by firms. Mandated public disclosure has expanded beyond the traditional accounting statements to include information on a firm's segmental performance, competitive environment, R&D expenditures, executive compensation, and pending litigation. Investors, however, also have access to a variety of costly private information. These sources range from analyst forecasts to illegally obtained "inside" information. Many studies have reported the impact of a public signal on traders' welfare absent other private information sources (e.g., Hakansson et al. 1982; Hirshleifer 1971; Marshall 1974; Ng 1975; Ohlson and Buckman 1981), and many others have reported on how private information acquisition will change in response to the public signal (e.g., Gonedes 1980; Gonedes et al. 1976; Hakansson 1977; Lundholm 1991; Verrecchia 1982a). However, little attention has been focused on how welfare will change in response to public information, taking into account both the direct effect of the public signal and its indirect effect due to the change in private information acquisition.

In this article, we present three different models designed to elucidate how the disclosure of public information can enhance the welfare of traders in a financial market when some or all traders also have access to costly private information. We begin our investigation by revisiting the work of Diamond (1985). In an economy where traders have the opportunity to acquire private signals about a single asset's payoff, Diamond shows that traders unanimously prefer the release of a public signal sufficiently precise to eliminate all private information acquisition. This is a potentially impor-

We would like to thank Anat Admati, Paul Pfleiderer, Mark Wolfson, and seminar participants at Stanford University for helpful discussions and comments.

Submitted April 1991.
Accepted June 1992.

tant result for accounting; positive welfare effects caused by public signals have been difficult to come by (see Verrecchia 1982a). We analyze the underlying cause of Diamond's result and its robustness to changes in the economy. Our first example shows that the result is not robust when the private signals have correlated rather than independent errors. However, our second example shows that Diamond's positive welfare result extends to an economy with multiple assets when the private signals have independent errors, regardless of the degree of correlation between the assets' underlying payoffs. The extension to an economy with multiple correlated assets is one way to assess the welfare properties of "information transfers."

Our third example considers an economy where some traders are allowed to acquire private information of varying quality while others are denied access to private information. Some traders clearly have differential access to information about a firm's prospects, possibly because of a unique relationship with the firm (e.g., an employee or supplier), or economies of scale in information acquisition and production (e.g., an analyst). We assess how a public signal can enhance the welfare of traders denied access to the information market, regardless of its effect on the informationally well-endowed traders. We believe that this final application reflects the regulatory concern that brought about the Security and Exchange Acts—that public information should be used to enhance the welfare of uninformed traders. We show, however, that the uninformed traders do not necessarily prefer to use a public signal to eliminate the informed traders' incentives to acquire information; they prefer this solution only when they are in the minority. Paradoxically, uninformed traders prefer to use public signals to enhance their welfare only when they may have insufficient power to do so. This occurs because, when the uninformed traders are in the majority, the loss in risk-sharing opportunities that accompany the public signal outweighs the benefit of informational parity.

Key Words: *Optimal disclosure, Private information acquisition, Social welfare.*

IN the following section, we derive an expression for ex *ante* expected utility in an economy with an arbitrary number of risky assets and an arbitrary private information structure. The assumptions regarding traders' preferences and the distribution of asset payoffs are such that, if the only source of information is a public signal, all traders prefer that the signal not be released; public information is a "bad." By conducting our analysis in an environment where public information in isolation has only negative welfare consequences, any positive welfare implications we find are entirely due to the interplay between the public and private information.

Before proceeding, two *caveats* are in order. First, all traders in the economies we consider have identical risk preferences, so optimal risk sharing would allocate an equal amount of the risky asset to each trader. Furthermore, this allocation could be accomplished by a round of trade prior to the arrival of any information. We preclude such a portfolio-balancing mechanism, and this is clearly a limitation to our analysis. In our defense, it is difficult to imagine how traders could credibly commit to forgo information acquisition prior to a "pre-information market." In Diamond's model and in our model, public information then becomes the lesser of two evils—although a public signal reduces welfare absent the availability of private information, it is not as bad as allowing the acquisition of private signals.

The second *caveat* is that we consider only an exchange economy. Consequently, each firm's payoff distribution (*ex ante*) is common knowledge; in a production economy, different firms have different payoff distributions. This creates a new demand for information from a social welfare point of view. That is, absent any information, all firms would have equal access to capital markets, yet this is clearly inefficient; adverse selection would arise and the need to mitigate it would create a new demand for public information signals.

I. Ex *Ante* Expected Utility

There are three periods in the model. In the first period traders are endowed with units of risky and riskless assets, choose the amount of costly information to acquire in the private information market, and then observe any publicly available information (these events can occur in any order). In the second period, units of the risky and riskless assets are traded in a speculative market, and asset payoffs are realized and consumption takes place in the third period. In the first period, traders know the precision that will be freely available by observing the assets' prices in the speculative market.[1] The speculative market consists of a single riskless asset whose value is always 1 at the end of the period and serves as the numeraire, and m risky assets with payoffs denoted by the random column vector $\tilde{F}$, which has covariance matrix H^{-1} (precision matrix H). The aggregate per capita supply of the risky assets is given by the random vector $\tilde{X}$, which has covariance matrix W.

There is a continuum of traders indexed by $v \in (0,1)$. Each trader v has the same utility function for terminal wealth, $U(w) = -\exp(-w/\rho)$, where ρ is the risk tolerance parameter. Each trader is endowed with B_v units of the riskless asset and X_v units of the risky assets (an m-dimensional vector). The endowments $\tilde{X}_v$ are mutually independent and, because there are an infinite number of traders, $\tilde{X}_v$ is also independent of $\tilde{X}$, the vector of per capita supplies of risky assets. For simplicity, let the vector $E(\tilde{X}_v) = 0$ for all v.

Each trader can purchase the private signal vector $\tilde{Y}_v$, which is correlated with $\tilde{F}$; a trader who does not purchase a private signal can be thought of as receiving a $\tilde{Y}_v$ that is uncorrelated with $\tilde{F}$. At this point there is no other structure placed on $\tilde{Y}_v$; it has an arbitrary variance matrix and covariance with $\tilde{F}$, and may differ across traders. The public signal vector $\tilde{Z}$ is observed by all, as is the risky asset price vector $\tilde{P}$. Finally, the vector $(\tilde{F}, \tilde{Y}_v, \tilde{Z}, \tilde{X}, \tilde{P})'$ is assumed to have a multivariate normal distribution, with nonsingular covariance matrix.

The fact that traders condition their beliefs on $\tilde{P}$, but the supply of the risky asset is stochastic, makes this a "noisy" rational expectations model. The purpose of the noise is to avoid a well-known nonexistence result in fully revealing rational expectations models. Specifically, in the absence of noise, price will fully reveal all private information but, if price is fully revealing, there is no incentive to acquire costly private information (see Grossman 1977; Grossman and Stiglitz 1980; Kreps 1977). With a stochastic per capita supply of the risky asset, price is still an informative variable, but it is not perfectly informative.[2]

---

[1] This is actually an *equilibrium* statement about the information market. In equilibrium, traders know the stochastic properties of the prices in the speculative market, which depend on the choices made in the information market.

Ex ante expected utility is evaluated after endowments are realized but before any signals are realized, and before trade takes place in the speculative market. Although all endowments are realized, a single trader observes only his or her own endowment; consequently, $\tilde{X}$ is still unknown and the price vector is still a noisy signal of the asset payoff vector. At this point in time, expected utility does not depend on any of the signal realizations, but only on their statistical properties, and hence it is an appropriate point to evaluate the welfare effects of different public signal disclosure policies.[3]

With negative exponential utility and normally distributed random variables, a trader's demand for risky assets is well-known (see Admati 1985). Demand is given by the $1 \times m$ vector,

$$D_{\upsilon} = \rho[E(\tilde{F}|Z, Y_{\upsilon}, P) - P]' \operatorname{Var}(\tilde{F}|Z, P, Y_{\upsilon})^{-1}. \tag{1}$$

By denoting the amount trader υ spends on private information acquisition as C_{υ}, his or her *ex ante* end-of-period wealth is given by:

$$\tilde{W}_{2\upsilon} = \rho[E(\tilde{F}|\tilde{Z}, \tilde{P}, \tilde{Y}_{\upsilon}) - \tilde{P}]' \operatorname{Var}(\tilde{F}|Z, P, Y_{\upsilon})^{-1}(\tilde{F} - \tilde{P}) + B_{\upsilon} - C_{\upsilon} + \tilde{P}' X_{\upsilon}.^{4} \tag{2}$$

It is the presence of the term $\tilde{P}' X_{\upsilon}$ that complicates the derivation of expected utility. Admati and Pfleiderer (1987) derive expected utility for traders endowed only with units of the riskless asset ($X_{\upsilon} = 0$ for all υ), and the stochastic supply of the risky asset is exogenous to the model (it is interpreted as arising from liquidity traders). For their purpose, examining the relative value of information, this assumption is without loss of generality. However, for the purpose of making welfare comparisons, the source of the risky asset supply needs to remain in the model. The following lemma derives ex *ante* expected utility in an economy with risky asset endownments (the proof is in the appendix).[5]

---

[2] An alternative method of introducing noise into the economy is to assume there is another class of traders whose demand is uncorrelated with any information, perhaps because of liquidity needs (see, e.g., Admati and Pfleiderer 1986). For welfare analysis, however, it is unappealing to have a group of traders whose wealth and preferences for wealth lie outside the model.

[3] This is the standard definition of *ex ante* expected utility. See Verrecchia (1982b), Diamond (1985), or Bushman (1991) for other examples. Also, if expected utility is evaluated prior to the endowment realization, the integral diverges for some parameter values.

[4] Although *ex ante* utility is computed prior to the realizations of $\tilde{Z}$, $\tilde{P}$, and $\tilde{Y}_{\upsilon}$, for normally distributed random variables, posterior variances do not depend on the signal realizations; they depend only on the parameters of the covariance matrix. Hence, $\operatorname{Var}(\tilde{F}|Z, P, Y_{\upsilon})$ is not a random variable and tildes will be left off the signals. Note, however, that $E(\tilde{F}|\tilde{Z}, \tilde{P}, \tilde{Y}_{\upsilon})$ does depend on the signal realizations.

[5] In the specific models that follow, the cost function is discrete and exogenous. Both Kim (1991) and Bushman (1991) examine the robustness of Diamond's (1985) result to changes in the way private information is acquired. Kim (1991) considers an economy where traders choose the level of precision and cost is a continuous, increasing function of the chosen precision. Bushman (1991) allows cost and private signal precision to be set strategically by an information monopolist.

Lemma 1. The *ex ante* expected utility of a trader who is endowed with X_v units of risky assets and who will observe the public signal $\tilde{Z}$, the asset price vector $\tilde{P}$, and pay C_v to observe $\tilde{Y}_v$ is given by:

$$E[-\exp(-\tilde{W}_{2v}/\rho)|X_v] = -(\alpha_v)^{1/2}\exp\left[\frac{1}{2}\left(\frac{X_v}{\rho}\right)'\beta\left(\frac{X_v}{\rho}\right)\right]$$

$$\exp\left[-\frac{B_v - C_v + E(\tilde{F})'X_v}{\rho}\right],$$

where:

$$\alpha_v = \frac{\det[\mathrm{Var}(\tilde{F}|Y_v, Z, P)]}{\det[\mathrm{Var}(\tilde{F}-\tilde{P})]},$$

$$\beta = \left\{(I_m I_m)\begin{bmatrix}\mathrm{Var}(\tilde{P}) & \mathrm{Cov}(\tilde{P},\tilde{F}) \\ \mathrm{Cov}(\tilde{P},\tilde{F})' & \mathrm{Var}(\tilde{F})\end{bmatrix}^{-1}(I_m I_m)'\right\}^{-1},$$

and I_m is an m-dimensional identity matrix. Also, $\alpha \in (0,1)$ and β is positive definite.

We will sometimes suppress the trader index, with the understanding that α varies with the precision of the private signal a particular trader receives. Note that utility is negative, so traders prefer it to be as close to zero as possible. In particular, they prefer α and

$$\left(\frac{X_v}{\rho}\right)'\beta\left(\frac{X_v}{\rho}\right)$$

to be as small as possible.

To begin to interpret the expected utility expression, note that the last term is the utility associated with the expected value of the trader's endowment. This value is then adjusted by α and β. We will refer to β as the *uncertainty of endowed wealth term* because it becomes more important in a trader's utility function as the trader's risky asset endowment vector gets uniformly larger, and does not enter the utility expression for a trader with no risky asset endowment. The positive welfare effect of a larger endowment shows up in the $E(\tilde{F})'X_v$ component of the last term in expected utility; β captures the negative welfare effect of the uncertainty of the endowment's value. Consistent with this interpretation of β is the fact that it becomes more important as ρ decreases (the trader becomes less risk-tolerant). One particular force that acts on β is *signal risk*. Public signals, or publicly observed informative prices, increase the uncertainty of endowed wealth because the holder of the endowment doesn't know if the signal will reveal the endowment as of high or low value.

We will refer to α as the *gains to speculative trade term* because it is the only adjustment to the utility of endowed wealth when the trader has no risky assets ($X_v=0$). With no endowment of the risky asset, traders can improve their expected welfare only by trading in the speculative market. More generally, because $\alpha \in (0,1)$, it represents the gain in utility from trading relative to the utility of the expected endowment of wealth (or, for the trader with $X_v=0$, the utility of endowed wealth).

An important special case of Lemma 1 is that of a single risky asset. Hellwig (1980), Verrecchia (1982b), and Bushman (1991), among others, have derived this case from first principles for particular information structures and prices. Simplifying Lemma 1 gives an expression for expected utility for an arbitrary information structure and price.

Corollary 1. When there is only one risky asset in the economy, *ex ante* expected utility is given by:

$$E[-\exp(-\tilde{W}_{2\upsilon}/\rho)|X_\upsilon]=-(\alpha_\upsilon)^{1/2}\exp\left[\frac{1}{2}\left(\frac{X_\upsilon}{\rho}\right)^2\beta\right]\exp\left[-\frac{B_\upsilon-C_\upsilon+E(\tilde{F})X_\upsilon}{\rho}\right],$$

where:

$$\alpha_\upsilon=\frac{\mathrm{Var}(\tilde{F}|Z,P,Y_\upsilon)}{\mathrm{Var}(\tilde{F}-\tilde{P})}\text{ and}$$

$$\beta=\frac{\mathrm{Var}(\tilde{P}|F)\mathrm{Var}(\tilde{F})}{\mathrm{Var}(\tilde{F}-\tilde{P})}.$$

The single-asset case is very useful for gaining intuition into the utility trade-offs induced by a public signal, and we will refer to Corollary 1 in two of the three sections that follow. At this point, however, we would like to stress the generality of Lemma 1. It applies for an arbitrary number of risky assets, which can have any type of cross-sectional covariance, and the public and private information structures are completely arbitrary. Of course, Lemma 1 is stated in terms of $\tilde{P}$, and we have not given a general characterization of the rational expectations price vector. The next lemma gives a rational expectations price and maintains much of the generality of Lemma 1, although it places more structure on the public and private information environment. The proof is a straightforward extension of Admati (1985) and is available from the authors on request.

Lemma 2. Suppose the private signal vector takes the form $\tilde{Y}_\upsilon=\tilde{F}+\tilde{\epsilon}_\upsilon$, where the stochastic process $(\epsilon_\upsilon)_{\upsilon\in(0,1)}$ is independent and each ϵ_υ has a zero mean vector and covariance matrix S_υ^{-1} (precision matrix S_υ). Also, the public signal vector takes the form $\tilde{Z}=\tilde{F}+\tilde{\epsilon}_p$, where ϵ_p is independent of all ϵ_υ, has a zero mean vector, and covariance matrix D^{-1} (precision matrix D). Along with $\mathrm{Var}(\tilde{F})=H^{-1}$ and $\mathrm{Var}(\tilde{X})=W$ given earlier, let

$$Q\equiv\rho\int_0^1 S_\upsilon d\upsilon.$$

For this information structure, the unique linear rational expectations price vector is given by,

$$\tilde{P}=A_0E(\tilde{F})+A_p\tilde{Z}+A_1\tilde{F}-A_2\tilde{X},$$

where:

$$A_0=\rho[\rho(H+D)+\rho QW^{-1}Q+Q]^{-1}H,$$
$$A_p=\rho[\rho(H+D)+\rho QW^{-1}Q+Q]^{-1}D,$$
$$A_1=[\rho(H+D)+\rho QW^{-1}Q+Q]^{-1}(Q+\rho QW^{-1}Q),\text{ and}$$
$$A_2=[\rho(H+D)+\rho QW^{-1}Q+Q]^{-1}(I_m+\rho QW^{-1}).$$

Note that the private information structure can still accommodate a wide variety of signals. As long as signals are independent across traders, they may be correlated across assets. A trader may be uninformed about some or all of the assets by making some of the diagonal elements in S_υ zero. Similarly, the public signals can be correlated in any way as long as the errors are independent of the private signal errors. The price in an economy without private signals is found by setting S_υ equal to the zero matrix for all υ (which implies that Q equals the zero matrix); the price in an economy without public signals is found by setting D equal to the zero matrix.[6]

Much of the anlaysis in the next three sections centers on the amount of information that prices convey to traders. It is useful to normalize the price vector so that it takes the form of other signals in the model. The following transformation accomplishes this:

$$\tilde{\boldsymbol{P}}=A_1^{-1}[\tilde{P}-A_0E(\tilde{F})]=\tilde{F}-A_1^{-1}A_2\tilde{X}=\tilde{F}-Q^{-1}\tilde{X}. \tag{3}$$

Because the transformation is linear, the normalized price $\tilde{\boldsymbol{P}}$ is informationally equivalent to the original price $\tilde{P}$. Thus, the precision gained from observing price is $\mathrm{Var}(\tilde{\boldsymbol{P}}|F)^{-1}$.

One final bit of machinery is needed. Traders in the following three examples choose whether to acquire costly private information. A trader's willingness to pay for private information, Φ, is the amount that equates the trader's utility without the private signal to the utility of paying Φ and observing the private signal. Replacing C_υ with Φ in Lemma 1 and solving for Φ gives the following lemma.

Lemma 3. A trader's willingness to pay for the private signal vector $\tilde{Y}_\upsilon$, given that $\tilde{Z}$ and $\tilde{P}$ are to be observed, is,[7]

$$\Phi=\frac{\rho}{2}\log\left\{\frac{\det[\mathrm{Var}(\tilde{F}|Y_\upsilon,Z,P)^{-1}]}{\det[\mathrm{Var}(\tilde{F}|Z,P)^{-1}]}\right\}.$$

II. Reversing Diamond's Result by Changing the Information Structure

As a point of comparison, we first present the positive welfare result of Diamond (1985) and then show that the result depends crucially on the structure of the private information. In particular, the private information signals in Diamond's model have independent errors. We show that if the signal errors are perfectly correlated, the welfare effect of releasing public information is no longer clear-cut.

Independent Private Signals

In Diamond's (1985) economy, there is a single risky asset and investors can pay c for a private information signal. Each private signal is of the form $\tilde{Y}_\upsilon=\tilde{F}+\tilde{\epsilon}_\upsilon$, where all the $\tilde{\epsilon}_v$ are independent and identically distributed with precision s. Let λ denote the fraction of traders who choose to purchase $\tilde{Y}_v$.

---

[6] Rational expectation prices for other information structures can be found in Lundholm (1988), who considers correlated public and private information, and Indjejikian (1991), who considers correlated private information.

[7] Admati and Pfleiderer (1987) also derive this expression. A trader's private information enters the expected utility expression only through $\mathrm{Var}(\tilde{F}|Z,P,Y_\upsilon)$ in α_υ. Consequently, when an agent compares the expected utility associated with different information acquisition decisions, all the exponential terms in Lemma 1 cancel, and the value of information is a function only of the relative values of $\mathrm{Var}(\tilde{F}|Z,P,Y_\upsilon)$ for each decision.

As the fraction of informed traders λ increases, $\tilde{P}$ becomes a more informative signal, and the willingness to pay for $\tilde{Y}_v$ decreases. Denote by λ^* the equilibrium value of λ that equates the willingness to pay Φ with the exogenous signal cost c (we are considering only interior solutions; $0<\lambda^*<1$). The willingness to pay for Y_v also decreases as d, the precision of the public signal, increases. To denote Φ's dependence on λ and d, we will write it as $\Phi(\lambda,d)$, so in equilibrium, $\Phi(\lambda^*,d)=c$. Denote by d^* the value of d that yields $\lambda^*=0$.[8] Diamond (1985) shows that traders' utility in an economy with a public signal of precision d^* and $\lambda^*=0$ (call this the *public signal economy*) is greater than their utility in an economy with $d=0$ and $0<\lambda^*<1$ (call this the *private signal economy*). By subscripting α and β with "pub" or "pri" to denote their values in the public or private signal economy, respectively, Diamond's results can be stated in terms of Lemma 1:

> *Proposition 1.* With independent private signals α_{pub} is always less than α_{pri}, and β_{pub} is always less than β_{pri}. Thus, traders unanimously prefer the public signal economy to the private one.

The next lemma helps to explain Diamond's (1985) results.

> *Lemma 4.* With independent private signals, the precision d^* that eliminates all private information acquisition in the public signal economy equals the precision gained from observing price in the private signal economy; that is, at $d^* = \text{Var}(\tilde{P}_{pri}|F)^{-1}$, $\lambda^*_{pub}=0$.

To see why changing to the public signal economy results in a welfare improvement, consider how the gains to speculative trade and the uncertainty of endowed wealth (α and β) vary across the two economies. With regard to β, Lemma 4 shows that the public signal imposes no additional signal risk on traders because it exactly replaces the signal risk that price created in the private signal economy. The public signal economy has the additional advantage of removing the uncertainty from the aggregate supply. In the public signal economy, the aggregate supply is known to traders who can infer $\tilde{X}$ by observing $\tilde{P}$ and $\tilde{Z}$, but it is unknown to traders in the private signal economy who observe only $\tilde{P}$ and $\tilde{Y}_v$. With regard to α, the informed traders in the private signal economy recoup their information cost c by gaining a lower α_v than the uninformed (recall that in equilibrium the utility of both groups is the same). That is, they enjoy higher gains from speculative trade than the uninformed. Consequently, the uninformed traders prefer the public signal economy because they are no longer trading at an informational disadvantage. Although informed traders lose their informational advantage in the public signal economy, they are no longer paying c for information.

The key assumption leading to Diamond's (1985) result is the independent private signal information structure. The sensitivity of the result to this assumption is apparent in an alternative information structure.

---

[8] We begin with economies in which, absent the public signal, $\lambda^* \in (0,1)$ and then find the value of d that causes λ^* to just meet the corner condition: the *minimum* value of d that causes $\lambda^*=0$.

Perfectly Correlated Private Signals

In this information structure, each trader who purchases a private signal sees the exact same signal: $\tilde{Y}_v=\tilde{F}+\tilde{\epsilon}$ for all v who purchase the signal. We cannot use Lemma 2 to find the price in this economy, but it is straightforward to derive the following price (see Grossman and Stiglitz 1980):

$$\tilde{P}=a_0E(\tilde{F})+a_1\tilde{Y}-a_2\tilde{X}, \tag{4}$$

where,

$$a_0=1-a_1,$$

$$a_1=\frac{\lambda s(w+\lambda\rho^2 s)}{h(w+\lambda^2\rho^2 s)+\lambda s(w+\lambda\rho^2 s)}, \text{ and}$$

$$a_2=\frac{w+\lambda\rho^2 s}{\rho h(w+\lambda^2\rho^2 s)+\rho\lambda s(w+\lambda\rho^2 s)}.$$

The next proposition shows that in the economy with perfectly correlated private signals, Diamond's (1985) unambiguous welfare result no longer holds.

Proposition 2. With perfectly correlated private signals, α_{pub} is always less than α_{pri}, but the ordering of β_{pub} and β_{pri} is generally ambiguous, so that, for some parameters, traders with sufficiently large endowments of the risky asset prefer the private signal economy while the remaining traders prefer the public signal economy.

The welfare ambiguity arises because $\beta_{pub}<\beta_{pri}$ if $(\lambda^*-2)w^2+h\rho^2\lambda w+\lambda^{*3}\rho^4 s(h+s)<0$; otherise $\beta_{pub}\geq\beta_{pri}$.[9] Thus, if $h=w=\rho=1$, $s=2$, and $c=\log(2)/2$, then $\lambda^*=0.408$ and $\beta_{pub}<\beta_{pri}$, so traders unanimously prefer the public signal economy. However, if we change ρ to 6, for instance, then $\lambda^*=0.266$ and $\beta_{pub}>\beta_{pri}$, so traders with large X_v prefer the private signal economy while traders with small X_v prefer the public signal economy (recall that the importance of β in a trader's utility depends directly on the size of X_v).

To prove the result, set S_v equal to zero in Lemma 2 to find the public signal economy price; the private signal economy price is given in equation (4). To find d^*, note that $\Phi(0,d^*)=c$ in the public signal economy and $\Phi(\lambda^*,0)=c$ in the private signal economy. Equate these two conditions and solve for d^* by using Lemma 3 to express Φ. Finally, insert the prices into Corollary 1 to find α_{pri} and β_{pri} in the private signal economy and α_{pub} and β_{pub} in the public signal economy.

---

[9] In the correlated private signal economy:

$$\lambda^*=0 \text{ if } e^{2c/\rho}-1\geq s/h,$$

$$\lambda^*=1 \text{ if } e^{2c/\rho}-1\leq\frac{ws}{s(s+h)\rho^2+wh}, \text{ and}$$

$$\lambda^*=\frac{\sqrt{w}}{\rho\sqrt{s(s+h)}}\sqrt{\frac{s}{e^{2c/\rho}-1}-h}$$

otherwise.

The source of ambiguity regarding the relative welfare of the public or private signal economy is directly attributable to the relative signal risk in the two economies, as the next lemma shows:

Lemma 5. In the economy with perfectly correlated private signals, the precision d^* required to eliminate private information acquisition in the public signal economy is greater than the precision available from observing price in the private signal economy.[10] That is, at $\lambda^*_{pub}=0, d^*>\operatorname{Var}(\tilde{P}_{pri}|F)^{-1}$.

To derive this result, equate $\Phi(\lambda^*,0)=c$ from the private signal economy with $\Phi(0,d^*)=c$ from the public signal one. Equating these two conditions and simplifying gives:

$$\frac{\rho}{2}\log\left[\frac{h+s}{h+\operatorname{Var}(\tilde{P}_{pri}|F)^{-1}}\right]=\frac{\rho}{2}\log\left[\frac{h+d^*+s}{h+d^*}\right].$$

At $d^*=\operatorname{Var}(\tilde{P}_{pri}|F)^{-1}$, the $LHS<RHS$. The result then follows because the RHS is decreasing in d.

Lemma 5, in contrast to Lemma 4, states that when the private signals are perfectly correlated, the public signal economy leaves the uninformed traders better informed than they were in the private one. The gains to speculative trade increase because the uninformed are no longer trading at an informational disadvantage (so $\alpha_{pub}<\alpha_{pri}$), but the increased precision can impose more signal risk on traders ($\beta_{pub}>\beta_{pri}$).[11]

III. An Extension of Diamond's Result to a Multiple Asset Economy

Although Diamond's (1985) result is not robust to changes in the error structure of private signals, it appears quite robust to a change in the underlying structure of the assets. Specifically, we consider an economy with two risky assets whose payoffs are correlated, possibly due to common industry membership, so that:

$$\operatorname{Var}\begin{bmatrix}\tilde{F}_1\\ \tilde{F}_2\end{bmatrix}=H^{-1}=\begin{bmatrix}h_1^{-1} & k\\ k & h_2^{-1}\end{bmatrix}.$$

Traders can pay c to acquire the private signal $\tilde{Y}_{1v}=\tilde{F}_1+\tilde{\epsilon}_{1v}$, where all the $\tilde{\epsilon}_{1v}$ are independent and identically distributed with precision s. In the public signal economy $\tilde{Z}=\tilde{F}_1+\tilde{\epsilon}_p$, where $\tilde{\epsilon}_p$ is independent of all $\tilde{\epsilon}_{1v}$ and has precision d. Thus, the first risky asset is exactly like the asset in Diamond's (1985) model, and the private signals and the public signal in the economy pertain (directly) only to that asset. There are no (direct) information signals available for the second asset. However, because the payoff of the second risky asset is correlated with the payoff of the first, the public and private signals provide information about the second asset's payoff as well.[12]

---

[10] In the correlated private signal economy, $\operatorname{Var}(\tilde{P}_{pri}|F)^{-1}$ is less than in the independent private signal economy because there are an infinite number of signals of precision s in the independent private signal economy but only one in the correlated private signal economy.

[11] Unlike the result for the independent private signal economy in Lemma 4, there is no obvious difference in traders' uncertainty regarding aggregate supply. An informed trader observing $\tilde{Y}$ and $\tilde{P}$ knows as much about the aggregate supply in the private signal economy as can be learned by observing $\tilde{Z}$ and $\tilde{P}$ in the public signal economy. However, uninformed traders learn more about the aggregate supply by observing $\tilde{Z}$ and $\tilde{P}$ in the public signal economy than by observing only $\tilde{P}$ in the private signal economy. The fact that β is smaller in the public signal economy than in the private one when λ^* is small is consistent with this interpretation.

The interpretation we have in mind is that a public announcement by one firm (say, of quarterly earnings) can affect traders' beliefs about nonannouncing firms in the same industry. By disclosing information of precision d*, the announcing firm may be improving the welfare of all traders endowed with its shares (as in Prop. 1). However, the indirect effect of this announcement may be to harm traders endowed with a nonannouncing firm's shares by imposing signal risk on them. Furthermore, because the private signal also provides indirect information about the nonannouncing firm, one might expect that the private signal's value will increase as the correlation between the assets increases. This might make it more difficult for the announcing firm to eliminate private information acquisition without imposing an excessive amount of signal risk on its shareholders.

Assume the per capita supplies of the risky assets are independent, so that:

$$\mathrm{Var}\begin{bmatrix}\tilde{X}_1\\ \tilde{X}_2\end{bmatrix}=W=\begin{bmatrix}w_1 & 0\\ 0 & w_2\end{bmatrix}.$$

Denote the fraction of traders who choose to acquire $\tilde{Y}_{1v}$ by λ. The price vector in the private signal economy is given in Lemma 2 by setting D equal to the zero matrix and noting that $Q=\rho\lambda S$, where:

$$S=\begin{bmatrix}s & 0\\ 0 & 0\end{bmatrix}.$$

Also, the incremental information gained by observing price is:

$$\mathrm{Var}(\tilde{P}_{pri}|F)^{-1}=QW^{-1}Q=\begin{bmatrix}(\rho\lambda s)^2/w_1 & 0\\ 0 & 0\end{bmatrix}.$$

A trader's willingness to pay for the private signal can be computed by using Lemma 3, and the rules for conditional distributions of normal random variables (which are given in the appendix) to find the necessary matrices. In particular,

$$\mathrm{Var}(\tilde{F}|P)^{-1}=H+QW^{-1}Q,$$

$$\det[\mathrm{Var}(\tilde{F}|P)^{-1}]=\frac{h_2[h_1+(\rho\lambda s)^2/w_1]}{1-k^2h_1h_2}, \tag{5}$$

$$\mathrm{Var}(\tilde{F}|Y_{1v},P)^{-1}=H+QW^{-1}Q+S^{-1}, \text{ and}$$

$$\det[\mathrm{Var}(\tilde{F}|Y_{1v},P)^{-1}]=\frac{h_2[h_1+s+(\rho\lambda s)^2/w_1]}{1-k^2h_1h_2}. \tag{6}$$

---

[12] For instance, the prior variance of $\tilde{F}_2$ is h_2^{-1}, but the posterior variance is 0,

$$\mathrm{Var}(\tilde{F}_2|Z)=h_2^{-1}-\frac{k^2dh_1}{h_1+d}.$$

Thus, the public signal provides more indirect information about the second asset as the asset covariance k increases.

The determinants given in equations (5) and (6) are known as generalized precisions, and the denominator of each is positive by the Schwarz inequality applied to H^{-1}. Both generalized precisions are increasing in k, the covariance in the asset payoffs, and the generalized precision of the informed trader increases faster than that of the uninformed. Thus, as the level of correlation between the assets increases, the informed traders reap larger precision gains than the uninformed. However, substituting the generalized precisions into the expression for a trader's willingness to pay for private information in Lemma 3 gives a surprising result; the denominators cancel out, and a trader's willingness to pay does not depend on the level of covariance between the risky assets. Specifically,

$$\Phi(\lambda^*,0)=\frac{\rho}{2}\log\left[\frac{h_1+(\lambda^*\rho s)^2/w_1+s}{h_1+(\lambda^*\rho s)^2/w_1}\right]$$

does not depend on k. The willingness to pay for private information depends on the gain in precision from being informed *relative* to being uninformed. Although the covariance in the asset payoffs increases the informed traders' precision more than that of the uninformed traders' precision, both levels of precision increase in exact proportion to those that would arise if the assets were uncorrelated ($k=0$). Hence, the relative gain in precision does not depend on k.

Because the willingness to pay for private information in this model is exactly the same as in Diamond's (1985) single asset model, the value of d^* that eliminates all private information acquisition is also the same, as seen in the following lemma.

> *Lemma 6*. In the two-asset economy with correlated payoffs, the precision d^* required to eliminate private information acquisition in the public signal economy equals the precision available from observing price in the private signal economy; that is, at $d^*=\mathrm{Var}(\tilde{P}_{pri}|F)^{-1}$, $\lambda^*_{pub}=0$.

The lemma is established by equating the equilibrium condition in the private signal economy with that in the public signal economy. That is, $\Phi(\lambda^*,0)=c=\Phi(0,d^*)$, or:

$$\frac{\rho}{2}\log\left[\frac{h_1+(\lambda^*\rho s)^2/w_1+s}{h_1+(\lambda^*\rho s)^2/w_1}\right]=\frac{\rho}{2}\log\left[\frac{h_1+d^*+s}{h_1+d^*}\right].$$

Solving for d^* gives $d^*=(\lambda^*\rho s)^2/w_1=\mathrm{Var}(\tilde{P}_{pri}|F)^{-1}$, exactly as in Lemma 4. The next proposition establishes the welfare result for this economy.

> *Proposition 3*. For any level of correlation between the risky asset payoffs, the public signal economy with public signal precision $d^*=\mathrm{Var}(\tilde{P}_{pri}|F)^{-1}$ is unanimously preferred to the private signal economy.

Some intermediate matrices necessary in the computation of α_{pri} and β_{pri} in the private signal economy are:

$$\mathrm{Var}(\tilde{F}-\tilde{P})=\mathrm{Var}[(I-A_1)\tilde{F}-A_2\tilde{X}]=(I-A_1)H^{-1}(I-A_1)'+A_2WA_2',\text{ and}$$

$$\begin{bmatrix}\mathrm{Var}(\tilde{P}) & \mathrm{Cov}(\tilde{P},\tilde{F})\\ \mathrm{Cov}(\tilde{P},\tilde{F})' & \mathrm{Var}(\tilde{F})\end{bmatrix}=\begin{bmatrix}A_1H^{-1}A_1'+A_2WA_2' & H^{-1}A_1'\\ A_1H^{-1} & H^{-1}\end{bmatrix}.$$

In the public signal economy, the price is found by setting Q equal to the zero matrix and

$$D=\begin{bmatrix} d^* & 0 \\ 0 & 0 \end{bmatrix}.$$

Some intermediate matrices necessary in the computation of α_{pub} and β_{pub} in the public signal economy are:

$$\mathrm{Var}(\tilde{F}-\tilde{P})=(H+D)^{-1}+\frac{1}{\rho^2}(H+D)^{-1}W(H+D)^{-1} \text{ and}$$

$$\begin{bmatrix} \mathrm{Var}(\tilde{P}) & \mathrm{Cov}(\tilde{P},\tilde{F}) \\ \mathrm{Cov}(\tilde{P},\tilde{F})' & \mathrm{Var}(\tilde{F}) \end{bmatrix}$$

$$=\begin{bmatrix} (H+D)^{-1}DH^{-1}+\frac{1}{\rho^2}(H+D)^{-1}W(H+D)^{-1} & H^{-1}D(H+D)^{-1} \\ (H+D)^{-1}DH^{-1} & H^{-1} \end{bmatrix}.$$

Using these expression in Lemma 1 to compute the α's and β's gives $\alpha_{pub}<\alpha_{pri}$ and $\beta_{pub}<\beta_{pri}$.

Proposition 3 extends Proposition 1 to the industry level. It shows that there are no externalities associated with public signal releases even when the firms in the economy have correlated payoffs. It remains socially optimal for individual firms to follow optimal disclosure policies without taking into account how the information transfer may affect other firms in the industry. Because the amount of public signal precision necessary to eliminate private information acquisition (d^*) does not depend on the correlation between the asset payoffs, no extra signal risk is imposed on traders in the public signal economy (relative to the private signal economy). Presumably this result could also be altered by changing the errors on the private signals from independent to correlated, as in section II. By analogy to Proposition 2, we expect that the uncertainty of endowed wealth term β would have an ambiguous sign. If so, we might also expect that traders with different endowments would prefer different economies, where endowments can now differ in both magnitude *and* composition across assets.

IV. Using Public Signals to Enhance the Welfare of Uninformed Traders

In this section, we turn our attention away from Diamond's model and study an economy with features of particular interest to regulators. In particular, an exogenous fraction of traders are endowed with the right to acquire costly private information—and they choose the level of precision they acquire—while the remaining traders are denied access to the information. The interpretation here is that a certain class of traders have access to information about a firm's prospects that other traders cannot acquire at any cost. The informed group potentially includes the firm's employees and registered insiders, suppliers and purchasers of the company's products, and the firm's financial anlaysts. The uninformed traders are denied access to information possibly because they have no direct relationship with the firm or possibly because economies of scale make information production infeasible (an argument that is put forth in Lev

1988). In this economy, we study how the use of public signals can further the welfare of the uninformed traders, without regard for the informed traders. Our aim is to capture the original spirit of the Security and Exchange Acts, which intended to "level the playing field" by mandating public information disclosure.

The nature of the costs facing the informed traders is an open question. In some sense they are endowed with information of a certain precision and a cost of acquisition arises only when additional precision is desired. However, illegal trading by registered insiders presumably bears some expected costs, and visits to company facilities by analysts certainly create some actual costs. The explicit modelling of the nature of these costs is beyond our purpose here. All we require is that the cost of private information to informed traders be increasing in its precision, with no restriction on how small the marginal cost may be.

As in the previously studied economies, the release of a public signal will reduce the private incentives to acquire information. Although eliminating private information acquisition leaves the informed and uninformed traders on equal footing, we show that the uninformed prefer this outcome only in certain circumstances.

Denote the fraction of informed traders by λ. Each informed trader has access to a private signal of the form $\tilde{Y}_v=\tilde{F}+\tilde{\epsilon}_v$, where $v\in(0,\lambda)$, $\mathrm{Var}(\tilde{\epsilon}_v)=s_v^{-1}$ and all $\tilde{\epsilon}_v$ are independent. Each informed trader chooses a level of precision s_v, which costs $C(s_v)=cs_v$. With independent private signals and a continuum of traders, each informed trader can acquire information without considering how this decision affects the information conveyed by price. We are interested in informed traders who are sophisticated enough to have access to the information market, yet not so large as to be able to influence price. Thus, all traders act as price-takers in the speculative market.[13]

Informed traders will acquire private signal precision until the marginal benefit equals the marginal cost; since all the informed traders have identical preferences for information, in equilibrium they will all acquire the same level of precision (this will be verified below). Denote by $\Phi(s,d)$ a trader's willingness to pay for private information of precision s when the public signal has precision d. A trader will choose s^* such that:

$$\frac{d\Phi}{ds}\Big|_{s^*}=c.$$

Using Lemma 3 to express $\Phi(s,d)$ and then differentiating gives:

$$s^*=\max\left\{0,s\,\Big|\,\frac{2c}{\rho}[h+d+\mathrm{Var}(\boldsymbol{P}_{pri}|F)^{-1}+s]=1\right\}. \tag{7}$$

An individual informed trader's information acquisition decision does not affect the information conveyed by price, so $\mathrm{Var}(\boldsymbol{P}_{pri}|F)^{-1}$ is treated as a constant by the individual. For this reason, all informed traders face the same optimization problem; consequently, they will all choose the same s^* in equilibrium. Noting that, for this model, Q in Lemma 2 equals $\rho\lambda s$, and using the normalized price in equation (3), we see that $\mathrm{Var}(\boldsymbol{P}_{pri}|F)^{-1}=(\rho\lambda s)^2/w$. Thus, in equilibrium,

---

[13] To be clear, the informed traders in our model take into account the information conveyed by the asset price, but they do not strategically anticipate how their demand for the risky asset will affect the asset's price. The behavior of stategic insiders is modeled in articles by Kyle (1985, 1989); for a list of applications, see Admati (1991).

$$s^*=\max\left\{0,s\,|\,\frac{2c}{\rho}[h+d+(\rho\lambda s)^2/w+s]=1\right\}.^{14} \tag{8}$$

The values of α_{pri} and β_{pri} can be computed by using Corollary 1 and the price equation in Lemma 2 (setting $D=0$ and remembering that s is now endogenous). From equation (8) we see that $s^*=0$ for a sufficiently large value of d. To find the minimum value of d that leads to informed traders choosing $s^*=0$, note that price is uninformative when $s^*=0$, so $\mathrm{Var}(\boldsymbol{P}_{pri}|F)^{-1}=0$ at this point. Therefore, setting $\mathrm{Var}(\boldsymbol{P}_{pri}|F)^{-1}=0$ and $s=0$ in equation (7), and solving for d^* gives:

$$d^*=\frac{\rho}{2c}-h.$$

It will be convenient, however, to express d^* in terms of the s^* that arises in the private signal economy. To do this, note that in the private signal economy:

$$s^*=\max\left\{0,s\,|\,\frac{2c}{\rho}[h+\mathrm{Var}(\boldsymbol{P}_{pri}|F)^{-1}+s]=1\right\},$$

which implies,

$$s^*=\max\left\{0,s\,|\,\mathrm{Var}(\boldsymbol{P}_{pri}|F)^{-1}+s=\frac{\rho}{2c}-h\right\}.$$

Thus, for positive s^* we can substitute d^* for the RHS of the previous equation to get:

$$d^*=s^*+\mathrm{Var}(\boldsymbol{P}_{pri}|F)^{-1}=s^*+(\ \lambda s^*)^2/w. \tag{9}$$

The values of α_{pub} and β_{pub} can be computed by using this value of d^*, and the price in Lemma 2 can be found by setting $Q=0$. Comparing α_{pri} and β_{pri} with α_{pub} and β_{pub} gives the following result.

> *Proposition 4.* For a fixed fraction λ of informed traders who acquire independent private signals of precision s^* at cost cs^* in the private signal economy, the uninformed traders unanimously prefer the public signal economy to the private one only if they are in the minority $(1-\lambda<1/2)$.[15]

Comparing the uninformed traders' α_{pri} with α_{pub} and β_{pri} with β_{pub} gives $\alpha_{pub}<\alpha_{pri}$ if $[(\rho\lambda s^*)^2+h(2\lambda-1)\rho^2s^*]w^4+(2\lambda s^*+h)\lambda^2\rho^4s^{*2}w^3+\lambda^4\rho^6s^{*4}w^2>0$, $\alpha_{pub}\geq\alpha_{pri}$ otherwise; and $\beta_{pub}<\beta_{pri}$ if $(2\lambda-1)\rho^4s^*w^2>0$, $\beta_{pub}\geq\beta_{pri}$ otherwise.

---

[14] Verrecchia (1982b) finds this equilibrium s^* for the economy with $\lambda=1$.

[15] Although quite tedious, it can also be shown that uninformed traders unanimously prefer *any* value of $d\in(0,d^*)$ only if they are in the minority, even though private information acquisition will persist for $d<d^*$. Kim (1991) presents a result similar to that of Proposition 4 when $\lambda=1$ (so there are no uninformed traders and the public signal economy is preferred), although he allows a more general cost function. He also sets the initial endowment so that α is the only relevant term in the utility expression.

It is difficult to imagine so many informed traders in the market that the uninformed traders are a minority; we typically think of the opposite extreme. If the firm's disclosure policy is set by the majority of shareholders, then Proposition 4 is particularly startling. It states that the only time the uninformed traders would prefer to use public signals to enhance their welfare is when they do not have sufficient voting power to do so.

To see why the uninformed traders are willing to tolerate their informational disadvantage for $\lambda < 1/2$, note from equation (9) that d^* provides the uninformed traders with s^* more precision in the public signal economy than they had in the private signal economy. Although the uninformed traders are no longer trading against informed traders with superior information in the public signal economy, the increased level of precision increases the signal risk they face. The uninformed are willing to tolerate their informational disadvantage in exchange for increased gains to trade as long as the fraction of informed traders is not too great.

Proposition 4 states the welfare effect for the uninformed traders. However, at the extreme values of λ, the informed and uninformed traders' preferences for a private or public signal economy are aligned.[16] When λ is near 0, there are few differences between uninformed and informed traders and a public signal imposes signal risk on all traders without offering much improvement in risk sharing, so both groups prefer the private signal economy. When λ is near 1, the informed traders have few uninformed traders to exploit; they trade mostly with themselves. The inefficient risk sharing that results causes all traders to prefer the public signal economy. Preferences diverge at $\lambda = 1/2$, where the uninformed traders prefer the public signal economy while the informed traders prefer the private signal economy.

V. Conclusions

It would be reckless to make strong welfare statements about the effect of public information disclosure on the basis of the stylized economies presented here. Our purpose is less ambitious. Although the three models are very structured, we believe they highlight some basic trade-offs in utility that arise when public signals are released in economies with private information acquisition. In particular, utility is the product of three terms: the utility of the expected value of endowed wealth, the gains from speculative trade, and the uncertainty of endowed wealth. Considered in isolation, the direct effect of public information is to reduce traders' welfare by reducing the gains to trade and increasing the uncertainty of endowed wealth (by creating signal risk). In the presence of costly private information, however, the public signal has an indirect effect as well; releasing public information reduces the value of private information. Diamond's (1985) model illustrates that it is possible for the welfare gains associated with eliminating private information acquisition to overcome the negative direct welfare effects of the public signal. Our analysis shows that this result is sensitive to the amount of signal risk the public signal imposes on traders. If the private signals have perfectly correlated errors, the amount of public signal precision necessary to eliminate private information acquisition creates too much signal risk for traders with large endowments of the risky asset (who are particularly sensitive to this aspect of utility). However, Diamond's model does extend to an economy with multiple assets. Signal risk is not transferred to holders of other risky assets in the economy, regardless of the level of correlation between the asset payoffs. In this model, optimal disclosure policy can be set at the firm level without considering externalities, even though there will be an information transfer from disclosing to nondisclosing firms.

---

[16] For the informed traders, $\alpha_{pub} < \alpha_{pri}$ if $[(2\lambda^2-1)s^2+h(\lambda^2+4\lambda-3)s+h^2(2\lambda-1)]v^5s\rho^2+[\lambda(\lambda^3+4\lambda^2-\lambda-2)s^2+h(4\lambda^3+\lambda^2-2\lambda-1)s+h^2(\lambda-1)(\lambda+1)]v^4\rho^4s^2+[\lambda^2(2\lambda^3+2\lambda^2-2\lambda-1)s+2h\lambda^2(\lambda-1)(\lambda+1)]+(\lambda-1)(\lambda+1)\lambda^4\ {}^8s^6v^2>0$, and $\alpha_{pub} \geq \alpha_{pri}$ otherwise.

Finally, we use the developed utility expression to analyze an economy where some traders can choose the amount of private signal precision to acquire and others are denied access to the information market. We show that the uninformed traders will prefer a public signal to eliminate their informational disadvantage only if they are in the minority. By resolving uncertainty prior to the opening of the speculative market, the public signal reduces the opportunities to share risk. As long as the number of informed traders is not too great, uninformed traders are willing to tolerate exploitation by informed traders in exchange for greater risk-sharing opportunities.

Appendix

Some general facts regarding matrix algebra and its application when evaluating the properties of multivariate normal random variables are useful. First, if A and B are two conformable and invertible matrices, with $A+B$ also non-singular, then:

$$(A+B)^{-1}=A^{-1}(A^{-1}+B^{-1})^{-1}B^{-1} \text{ and}$$
$$(A+B)^{-1}=A^{-1}-A^{-1}(A^{-1}+B^{-1})^{-1}A^{-1}.$$

Second, let A, B, B', and D compose the submatrices of a partitioned matrix. The inverse of this partitioned matrix is given by:

$$\begin{bmatrix} A & B \\ B' & D \end{bmatrix}^{-1}=\begin{bmatrix} A^{-1}+FEF' & -FE \\ -EF' & E \end{bmatrix}.$$

where $E=[D-B'A^{-1}B]^{-1}$ and $F=A^{-1}B$.

Third, let X and Y be $m\times 1$ multivariate normal random vectors. Denote by $\mathrm{Cov}(X,Y)$ the matrix of element-by-element covariances between X and Y and $\mathrm{Var}(X)=\mathrm{Cov}(X,X)$. If Z is also an $m\times 1$ random vector then $\mathrm{Cov}(X,Y+Z)=\mathrm{Cov}(X,Z)+\mathrm{Cov}(X,Y)$. If A and B are matrices then $\mathrm{Cov}(AX,BY)=A\,\mathrm{Cov}(X,Y)B'$. Further,

$$\mathrm{Var}[X-E(X|Y)]=\mathrm{Var}(X|Y), \tag{A1}$$

$$\mathrm{Cov}[Y,X-E(X|Y)]=0, \text{ and} \tag{A2}$$

$$\mathrm{Cov}[X,X-E(X|Y)]=\mathrm{Var}(X|Y). \tag{A3}$$

Proof of Lemma 1

The proof extends the results of Admati and Pfleiderer (1987, 107–09) to the case of stochastic endowments of the risky asset. We express expected utility using iterated expectations, then use Admati and Pfleiderer's result to evaluate an inside expectation, taking $\tilde{P}$ as given. We then evaluate the outside expectation by identifying a quadratic form in $\tilde{P}$ and evaluating it with the moment-generating function of a normal random variable. The derivation is slightly different for agents who observe only price and those who observe exogenous information as well as price.

Lemma 1, Part I: Informed Agents

Consider the agent who observes some information ($\tilde{Y}_v$ and/or $\tilde{Z}$) in addition to price. Denoting $\mathrm{Var}(\tilde{F}|Z,P,Y_v)^{-1}$ by $\hat{R}$, the end-of-period wealth of the agent is given by:

$$\tilde{W}_{2v}=\rho[E(\tilde{F}|\tilde{Z},\tilde{P},\tilde{Y}_v)-\tilde{P}]'\hat{R}(\tilde{F}-\tilde{P})+B_v-C_o+\tilde{P}'X_v.$$

Now rewrite the above expression in terms of two new random variables that both have zero means when conditioned on $\tilde{P}$. Letting $\hat{\mu}=E(\tilde{F}-\tilde{P}|\tilde{P})$, define $\tilde{\xi}=\tilde{F}-\tilde{P}-\hat{\mu}=\tilde{F}-E(\tilde{F}|\tilde{P})$, and $\tilde{\eta}=\tilde{F}-E(\tilde{F}|\tilde{Z},\tilde{P},\tilde{Y}_v)$. Thus,

$$\tilde{W}_{2v}=\rho(\hat{\mu}+\tilde{\xi}-\tilde{\eta})'\hat{R}(\tilde{\xi}+\hat{\mu})+B_v-C_v+\tilde{P}X_v.$$

Note that price still appears as a random variable in final wealth. Conditioning it out, by the laws of iterated expectations, reduces the problem to one involving only $\tilde{\xi}$ and $\tilde{\eta}$, as in Admati and Pfleiderer (1987). To do this, write *ex ante* expected utility as:

$$\begin{aligned} U &= E_{\xi\eta p}[-\exp(-\tilde{W}_{2\nu}/\rho)] \\ &= E_P\{E_{\xi\eta}[-\exp(-\tilde{W}_{2\nu}/\rho)]\}. \end{aligned}$$

Consider the inside expectation and note that P is treated as a constant in $E_{\xi\eta}$. Writing the inside expectation as an integral gives:

$$E_{\xi\eta} = -\exp\left(-\frac{B_\nu - C_\nu + \tilde{P}X_\nu}{\rho}\right)\left(\frac{1}{(2\pi)^m|\hat{\Sigma}|^{1/2}}\right)\int_{R^{2m}}$$

$$\exp\left[-(\hat{\mu}+\xi-\eta)'\hat{R}(\hat{\mu}+\xi)-\frac{1}{2}(\xi',\eta')\,\hat{\Sigma}^{-1}\begin{pmatrix}\xi\\ \eta\end{pmatrix}\right]d(\xi'\eta'),$$

where,

$$\hat{\Sigma} = \mathrm{Var}(\tilde{\xi}',\tilde{\eta}'|P) = \mathrm{Var}(\tilde{\xi}',\tilde{\eta}')$$

since, by construction, $\tilde{P}$ is uncorrelated with $(\tilde{\xi}',\tilde{\eta}')$; (see eq. A2). Substituting in the definitions of ξ and η gives:

$$\hat{\Sigma} = \begin{pmatrix} \mathrm{Var}[\tilde{F}-E(\tilde{F}|\tilde{P})] & \mathrm{Cov}[\tilde{F}-E(\tilde{F}|\tilde{P}),\tilde{F}-E(\tilde{F}|\tilde{P},\tilde{Z},\tilde{Y}_\nu)] \\ \mathrm{Cov}[\tilde{F}-E(\tilde{F}|\tilde{P}),\tilde{F}-E(\tilde{F}|\tilde{P},\tilde{Z},\tilde{Y}_\nu)] & \mathrm{Var}[\tilde{F}-E(\tilde{F}|\tilde{P},\tilde{Z},\tilde{Y}_\nu)] \end{pmatrix}$$

$$= \begin{pmatrix} \hat{\Psi} & \hat{R}^{-1} \\ \hat{R}^{-1} & \hat{R}^{-1} \end{pmatrix} \text{where } \hat{\Psi} = \mathrm{Var}[\tilde{F}-E(\tilde{F}|\tilde{P})] = \mathrm{Var}(\tilde{F}|P)] = \mathrm{Var}(\tilde{F}-\tilde{P}|P).$$

The simplification of the diagonal terms in $\hat{\Sigma}$ follow from equation (A1) and the definitions of ξ and η. The simplification of the off-diagonal terms follows from equations (A1) and (A2), with the realization that $E(\tilde{F}|\tilde{P})$ and $\tilde{F}-E(\tilde{F}|\tilde{P},\tilde{Z},\tilde{Y}_\nu)$ are uncorrelated because $E(\tilde{F}|\tilde{P})$ is a linear function of $\tilde{P}$.

At this point we can rely on Admati and Pfleiderer's (1987) proof to evaluate $E_{\xi\eta}$. In particular, $\tilde{\xi}$ and $\tilde{\eta}$ have the same means as in Admati and Pfleiderer, and their covariance matrix $\hat{\Sigma}$ has the same structure as the covariance matrix Σ in Admati and Pfleiderer. Substitute $\hat{\mu}, \hat{\Psi}, \hat{\Sigma}$, and $\hat{R}$ for μ, Ψ, Σ, and R in equation (A1) of Admati and Pfleiderer, and, from their equation (A14), we get:

$$E_{\xi\eta} = -\exp\left(-\frac{B_\nu - C_\nu + \tilde{P}X_\nu}{\rho}\right)\frac{1}{|\hat{R}|^{1/2}|\hat{\Psi}|^{1/2}}\exp\left(-\frac{1}{2}\hat{\mu}'\hat{\Psi}^{-1}\hat{\mu}\right). \tag{A4}$$

Now consider the outside expectation. The only random vector in $E_{\xi\eta}$ is $\tilde{P}$. Writing out $E_P(E_{\xi\eta})$ as an integral gives:

$$U = -\exp\left(-\frac{B_\nu - C_\nu}{\rho}\right)\frac{1}{|\hat{R}|^{1/2}|\hat{\Psi}|^{1/2}(2\pi)^{m/2}|\mathrm{Var}(\tilde{P})|^{1/2}}\int_{R^m}\exp\left(-\frac{PX_\nu}{\rho}-\frac{\Xi}{2}\right)dP$$

where,

$$\Xi = [\tilde{P}-E(\tilde{F}|\tilde{P})]'\mathrm{Var}(\tilde{F}|P)^{-1}[\tilde{P}-E(\tilde{F}|\tilde{P})]+[\tilde{P}-E(\tilde{P})]'\mathrm{Var}(\tilde{P})^{-1}[\tilde{P}-E(\tilde{P})].$$

Using the fact that $E(\tilde{F}|\tilde{P}) = E(\tilde{F})+\mathrm{Cov}(\tilde{F},\tilde{P})'\mathrm{Var}(\tilde{P})^{-1}[\tilde{P}-E(\tilde{P})]$, and $E(\tilde{P}) = E(\tilde{F})$ since $E(X_\nu)=0$, one may write $\tilde{P}-E(\tilde{F}|\tilde{P}) = [I_m-\mathrm{Cov}(\tilde{F},\tilde{P})'\mathrm{Var}(\tilde{P})^{-1}][\tilde{P}-E(\tilde{P})]$, so that:

$$\Xi = [\tilde{P}-E(\tilde{P})]'\beta^{-1}[\tilde{P}-E(\tilde{P})]$$

where,

$$\beta^{-1} = [I_m-\mathrm{Cov}(\tilde{F},\tilde{P})'\mathrm{Var}(\tilde{P})^{-1}]'\mathrm{Var}(\tilde{F}|\tilde{P})^{-1}[I_m-\mathrm{Cov}(\tilde{F},\tilde{P})'\mathrm{Var}(\tilde{P})^{-1}]+\mathrm{Var}(\tilde{P})^{-1}.$$

Having converted Ξ into a quadratic form in $\tilde{P}$, the simplification of U is straightforward: the expression is the moment-generating function of a random variable with covariance matrix β^{-1}, and can be rewritten accordingly. Thus,

$$U=-\exp\left(-\frac{B_\nu-C_\nu}{\rho}\right)\frac{1}{|\hat{R}|^{1/2}|\hat{\Psi}|^{1/2}(2\pi)^{m/2}|\mathrm{Var}(\tilde{P})|^{1/2}}\int_{R^m}\exp\left\{-\frac{PX_\nu}{\rho}-\frac{1}{2}[P-E(P)]'\beta^{-1}[P-E(P)]\right\}dP$$

$$=-\frac{1}{|\hat{R}|^{1/2}|\hat{\Psi}|^{1/2}|\beta^{-1}|^{1/2}|\mathrm{Var}(\tilde{P})|^{1/2}}\exp\left[-\frac{B_\nu-C_\nu+E(\tilde{P})X_\nu}{\rho}+\frac{1}{2}\left(\frac{X_\nu}{\rho}\right)'\beta\left(\frac{X_\nu}{\rho}\right)\right].$$

$E(\tilde{P})$ may be replaced by $E(\tilde{F})$, and the rules for taking the inverse of a partitioned matrix indicate that β^{-1} is the sum of the components of:

$$\begin{bmatrix} \mathrm{Var}(\tilde{P}) & \mathrm{Cov}(\tilde{P},\tilde{F}) \\ \mathrm{Cov}(\tilde{P},\tilde{F})' & \mathrm{Var}(\tilde{F}) \end{bmatrix}^{-1}.$$

Finally, note that by setting $\tilde{X}_\nu \equiv 0$, the expression for utility reduces to the expression with nonstochastic initial wealth. Setting $\tilde{X}_\nu \equiv 0$ in our expression and equating the result with Admati and Pfleiderer's (1987) Proposition 3.1 shows that;

$$|\hat{\Psi}||\beta^{-1}||\mathrm{Var}(\tilde{P})|=|\mathrm{Var}(\tilde{F}-\tilde{P})|.$$

Substituting this into the previous expression for U establishes the result for traders who observe exogenous information and price.

Lemma 1, Part II: Uninformed Agents

For traders who observe only $\tilde{P}$ and no exogenous information, $\tilde{\xi}=\tilde{\eta}=\tilde{F}-E(\tilde{F}|\tilde{P})$. The previous proof is no longer valid because $\hat{\Sigma}$ is now singular. However, we can proceed in similar fashion by using only $\tilde{\xi}$. Specifically, for an agent who observes only price, $E(\tilde{\xi})=0$, $\mathrm{Var}(\tilde{\xi})=\mathrm{Var}(\tilde{F}|P)=\hat{R}$, and $\tilde{W}_{2\nu}=\rho\hat{\mu}'\hat{R}(\tilde{\xi}+\hat{\mu})+B_\nu-C_\nu+\tilde{P}X_\nu$. This gives:

$$U=E_P\{E_\xi[-\exp(-\tilde{W}_{2\nu}/\rho)]\}.$$

Considering the inside expectation gives:

$$E_\xi=-\exp\left(-\frac{B_\nu+\tilde{P}X_\nu}{\rho}-\hat{\mu}'\hat{R}\hat{\mu}\right)E_\xi[\exp(-\hat{\mu}'\hat{R}\tilde{\xi})],$$

Using the moment-generating function of $\tilde{\xi}$ gives:

$$E_\xi=-\exp\left(-\frac{B_\nu+\tilde{P}X_\nu}{\rho}-\hat{\mu}'\hat{R}\hat{\mu}\right)\exp\left(\frac{1}{2}\hat{\mu}'\hat{R}\hat{R}^{-1}\hat{R}'\hat{\mu}\right)$$

$$=-\exp\left(-\frac{B_\nu+\tilde{P}X_\nu}{\rho}\right)\exp\left(-\frac{1}{2}\hat{\mu}'\hat{R}^{-1}\hat{\mu}\right).$$

Because $\hat{R}=\hat{\Psi}^{-1}$ in this case, this expression equals equation (A4). The results of the previous proof following equation (A4) establishes the result for traders who observe only $\tilde{P}$.

Lemma 1, Part III: Characteristics of α and β

Finally, we show that $\alpha \in (0,1)$ and β is positive definite. First, $\alpha<1$ because $\mathrm{Var}(\tilde{F}|Z,P,Y_\nu)$ can be rewritten as $\mathrm{Var}(\tilde{F}-\tilde{P}|Z,P,Y_\nu)$ and, for normal random variables, the determinant of the conditional covariance matrix is less than the determinant of the unconditional covariance matrix; thus, $|\mathrm{Var}(\tilde{F}-\tilde{P}|Z,P,Y_\nu)| < |\mathrm{Var}(\tilde{F}-\tilde{P})|$. Second, $\alpha>0$ because the covariance matrix for any vector of normal random variables is positive definite, and the determinant of a positive definite matrix is positive.

Now consider β. The covariance matrix for the combined vector $(\tilde{P},\tilde{F})$ is;

$$\begin{bmatrix} \mathrm{Var}(\tilde{P}) & \mathrm{Cov}(\tilde{P},\tilde{F}) \\ \mathrm{Cov}(\tilde{P},\tilde{F})' & \mathrm{Var}(\tilde{F}) \end{bmatrix},$$

which is positive definite. Denote its inverse by C (a $2m \times 2m$ matrix) and note that the inverse of a positive definite matrix is positive definite. It follows from the definition of a positive definite matrix that, for all $1 \times 2m$ vectors and $X \neq 0$, $XCX'>0$. Since this holds for all $X \neq 0$ then it holds for $X=Y(I_m I_m)$ and all $1 \times mY$ vectors, where $Y \neq 0$. Thus, $Y(I_m I_m)C(I_m I_m)'Y'>0$, which establishes the result.

References

Admati, A. R. 1985. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica* 53 (May): 629–57.

———. 1991. The informational role of prices. *Journal of Monetary Economics* 28 (Autumn): 347–61.

———, and P. Pfleiderer. 1986. A monopolistic market for information. *Journal of Economic Theory* 39 (August): 400–38.

———, and ———. 1987. Viable allocations of information in financial markets. *Journal of Economic Theory* 43 (October): 76–115.

Bushman, R. 1991. Public disclosure and the structure of private information markets. *Journal of Accounting Research* 29 (Autumn): 261–76.

Diamond, D. W. 1985. Optimal release of information by firms. *Journal of Finance* 40 (September): 1071–94.

Gonedes, N. 1980. Public disclosure rules, private information-production decisions, and capital market equilibrium. *Journal of Accounting Research* 18 (Spring): 441–76.

———, N. Dopuch, and S. Penman. 1976. Disclosure rules, information production, and capital market equilibrium: The case of forecast disclosure rules. *Journal of Accounting Research* 14 (Autumn): 89–137.

Grossman, N. 1977. The existence of futures markets, noisy rational expectations and information externalities. *Review of Economic Studies* 14 (Autumn): 431–49.

———, and J. Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70 (June): 393–408.

Hakansson, N. 1977. Interim disclosure and forecasts: An economic analysis and framework for choice. *The Accounting Review* 52 (April): 396–416.

———, G. Kunkel, and J. Ohlson. 1982. Sufficient and necessary conditions for information to have social value in pure exchange. *Journal of Finance* 37 (December): 1169–81.

Hirshleifer, J. 1971. The private and social value of information and the reward to inventive activity. *American Economic Review* 61 (September): 561–74.

Hellwig, M. F. 1980. On the aggregation of information in competitive markets. *Journal of Economic Theory* 22 (June): 477–98.

Indjejikian, R. 1991. The impact of costly information interpretation on firm disclosure decisions. *Journal of Accounting Research* 29 (Autumn): 277–301.

Kim, O. 1991. Disagreements among shareholders and managers over firm's disclosure policy. *Journal of Finance* (Forthcoming).

Kreps, D. 1977. A note on "fulfilled expectations" equilibria. *Journal of Economic Theory* 14 (June): 32–43.

Kyle, A. 1985. Continuous auctions and insider trading. *Econometrica* 53 (November): 1315–35.

———. 1989. Informed speculation with imperfect competition. *Review of Economic Studies* 56 (Autumn): 317–56.

Lev, B. 1988. Toward a theory of equitable and efficient accounting policy. *The Accounting Review* 63 (January): 1–23.

Lundholm, R. 1988. Price-signal relations in the presence of correlated public and private information. *Journal of Accounting Research* 26 (Spring): 107–18.

———. 1991. Public signals and the equilibrium allocation of private information. *Journal of Accounting Research* 29 (Autumn): 322–49.

Marshall, J. 1974. Private incentives and public information. *The American Economic Review* 64 (June): 373–90.

Ng, D. 1975. Information accuracy and social welfare under homogeneous beliefs. *Journal of Financial Economics* 2 (March): 53–70.

Ohlson, J., and G. Buckman. 1981. Towards a theory of financial accounting: Welfare and public information. *Journal of Accounting Research* 19 (Autumn): 399–433.

Verrecchia, R. 1982a. Mathematical models in financial accounting. *Journal of Accounting Research* 20 (Supplement): 1–42.

———. 1982b. Information acquisition in a noisy rational expectations economy. *Econometrica* 50 (November): 1415–30.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 113–133

Auditee Incentives for Auditor Independence: The Case of Nonaudit Services

Mohinder Parkash
Oakland University
Carol F. Venable
San Diego State University

SYNOPSIS AND INTRODUCTION: The purpose of this research is to test the effects of agency incentives and knowledge spillovers in joint engagements for audit and nonaudit services. External parties and regulators have associated joint nonaudit purchases with beliefs about impaired auditor independence. Since reductions in perceived independence correspondingly reduce audit credibility, agency theory suggests that auditees have an incentive to limit nonaudit purchases because of the agency costs imposed when an audit's monitoring value is diminished.

The frequency of nonaudit services is expected to reflect knowledge spillovers from joint engagements and the level of economic bonding between the auditor and the auditee. Hence, two categories of services are examined, recurring and nonrecurring. Agency incentives are expected to significantly influence the auditee's decision to purchase recurring nonaudit services since the auditor's revenue from a recurring service can be perceived as an annuity. However, for a nonrecurring service, future knowledge spillovers and incremental economic bonding are expected to be minimal. Because the auditor's revenue from the nonrecurring service is not an additional annuity, any perceived impairment to independence and credibility caused by incremental economic bonding should be limited.

Data from *Fortune 500* companies, disclosed under Accounting Series Release (ASR) No. 250, are used in multiple regressions to test whether auditees vary nonaudit purchases in accordance with the potential agency

The authors wish to acknowledge the useful comments on earlier drafts of this paper received from Bob Capettini, Chee Chow, Gary Grudnitski, Bob Hagerman, Susan Moyer, Mike Shields, Ira Solomon, Earl Wilson, an anonymous reviewer and associate editor, and workshop participants at San Diego State University, George Mason University and Oakland University. The assistance of David Buck and Joanne Healy in partial data collection is acknowledged.

Submitted December 1989.
Accepted April 1992.

costs that could arise. The findings indicate that variables related to expected agency costs (management ownership, outside investment concentration, and leverage) significantly explain cross-sectional differences in the demand for recurring nonaudit services. The data also indicate that auditees purchase higher levels of recurring services when they engage particular auditors (i.e., industry specialists or the audit firm that is the largest overall provider of nonaudit services). Consistent with expectations, overall agency costs do not explain the level of nonrecurring services purchased from the auditor. This latter finding suggests that selection of an industry specialist may be a quality signal that allows an auditee to purchase higher levels of recurring nonaudit services than would normally be expected.

Future research may explore how nonaudit services influence auditor switching. For example, environmental factors could increase the need for nonaudit services, thus making it more efficient to change to an industry specialist. Alternatively, a reorganization of ownership or capital structure could reduce the need for services, which would lead to a renegotiation of an existing joint services contract. At a minimum, the results reported here suggest that auditees evaluate a potential auditor's ability to provide a particular type or level of nonaudit services.

Key Words: *Nonaudit services, Auditor independence, Agency costs, Economic bonding, ASR No. 250.*

Data Availability: *A list of sample firms is available from the first author upon request. Data used in this study were obtained from public sources that are identified in the text.*

IN the first section, we provide background information, review prior research, and develop separate hypotheses for two categories of nonaudit services, recurring and nonrecurring. The method used to test the hypotheses is presented in section II, and section III contains the results. The last section provides a summary and concluding remarks.

I. Agency Costs and the Procurement of Jointly Produced Services

For over a decade, various groups, including congressional committees, the Securities and Exchange Commission (SEC), the Public Oversight Board (POB 1986), and the Commission on Auditors' Responsibilities (CAR 1978), have asserted that auditor independence can be impaired when joint nonaudit services are provided.[1]

---

[1] There is no precise definition for the term "auditor independence" (Antle 1984). It has been viewed as a personal trait encompassing objectivity and impartiality. As such, it is neither directly observable nor readily assessable. Another view focuses on appearances, i.e., observable actions, events, and structures that are presumed to be related to objectivity and impartiality. These observable phenomena can arise *ex ante* or *ex post* to the audit process. *Ex ante*, they suggest a causal relation between the observation and the independence state that ensues; *ex post*, they imply an evidence relation between the observation and independence state that occurred. In this article, independence, in any of its forms, is assumed to be present in varying degrees that signify its extent rather than its absolute presence or absence.

Although no direct evidence of impairment has been found, controversy over nonaudit services continues.[2]

Beck et al. (1988a, 1988b) demonstrate, analytically and empirically, why rational observers may perceive an impairment to auditor independence. In their model of the supply side of the market for joint audit and nonaudit services, they identify conditions that give rise to economic bonding. Additionally, they describe the auditor's motivation to provide fee incentives to auditees for joint services. This research extends Beck et al. (1988a) to examine whether auditees vary nonaudit purchases in accordance with the potential magnitude of the costs that may be incurred with perceptions of impaired auditor independence.

Agency literature (Fama and Jensen 1983a, 1983b; Jensen and Meckling 1976; Watts 1977; Watts and Zimmerman 1983) suggests that auditees are sensitive to perceptions of impaired independence since credible auditors are used to reducing the principal-agent conflict. Specifically, agency theory posits that the demand for independent audits occurs because managers (auditees) want certification of their financial disclosures and verification of their actions. Absent independent monitoring, agency costs are imposed on auditees.[3] Since a perceived impairment to independence reduces the expected monitoring value of an audit, auditees are motivated to avoid such situations and to manage the perceived level of economic bonding between the auditor and themselves by controlling the proportion of joint audit and nonaudit services.

Prior Research on Nonaudit Services

In general, prior studies fall into two categories. The first category typically concentrates on the audit-pricing decision and uses data collected from direct survey instruments (Abdel-khalik 1990; Palmrose 1986a, 1986b, 1988; Simunic 1984). These studies provide conflicting results on whether nonaudit services have a bearing on audit fees, possibly because of inconsistencies in the definition of nonaudit services and response bias attributable to data sensitivity.[4]

The second category of studies attempts to overcome these limitations by utilizing public disclosures required by ASR No. 250 during the period 1978–81 (Beck et al. 1988b; DeBerg et al. 1991; Glezen and Millar 1985; Scheiner 1984; Scheiner and Kiger 1982). In general, these studies provide descriptive results and limited tests of the association between nonaudit services and various independence-related measures. They do not, however, assess whether auditees monitor and control the level of nonaudit services purchased from their auditor.

---

[2] Although various aspects of the relation between nonaudit services and auditor independence were discussed by the SEC in ASR Nos. 250, 296, and 304 (SEC 1978, 1981, 1982), formal SEC investigations and conflicting opinions about the appropriateness of nonaudit services continue to occur (Berton 1989; Mednick 1990).

[3] According to Jensen and Meckling (1976, 308), agency costs include: (1) monitoring expenditures, (2) bonding expenditures, and (3) the residual loss. In this study, agency costs may include one or more of these components.

[4] All of these studies have relatively low response rates (20 to 33 percent) which suggests a possible nonresponse bias attributable to the sensitive nature of the requested disclosures. Further, while Simunic (1984) specifically excluded tax-related work in defining nonaudit services, Palmrose (1986b) included it and found varying results when auditees were categorized by size. Abdel-khalik (1990), utilizing management advisory services data, found results that differed from both Simunic and Palmrose, but it is not clear how survey respondents interpreted the terminology (i.e., management advisory services) and types of services.

Glezen and Millar (1985, 869) examine stockholder voting behavior before and after mandated disclosure of nonaudit services and find no change in auditor approval ratios. Although they conclude "that the levels of nonaudit services for the firms studied were not perceived by stockholders to affect independence," an expanded interpretation of their findings suggests that auditees may have anticipated the potential impairment of their auditor's credibility and accordingly restricted their purchases of nonaudit services. Anecdotal evidence to support this contention is contained in comment letters to the SEC, and some limited empirical support comes from DeBerg et al. (1991), who find that companies reduce the level of recurring nonaudit services when they switch auditors. This classification of nonaudit services as recurring and nonrecurring was initially identified by Beck et al. (1988a, 1988b) in developing and testing an analytical model of the supply side of a competitive market for joint audit and nonaudit services.

The work by Beck et al. (1988a, 1988b) is pertinent in examining the relation between agency incentives and nonaudit services because it suggests that nonaudit services result in incremental economic bonding beyond what DeAngelo (1981) associated with an individual audit engagement. Their model further suggests that part of the demand for certain nonaudit services is attributable to knowledge spillovers and auditor-provided economic incentives (e.g., fee discounts that represent cost savings from knowledge spillovers in joint engagements). In a competitive market, Beck et al. suggest that auditors use a "limit" pricing strategy and offer fee discounts so auditees would not obtain nonaudit services from competitors. The auditor's motivation is to limit competitors' access to information on cost functions that would improve their ability to bid on future audit engagements. Beck et al. (1988a, 1988b) report consistent empirical evidence of a positive association between recurring nonaudit services and the length of the auditor's tenure.

Recurring Nonaudit Services

The analytical model in Beck et al. (1988a) demonstrates that start-up and switching costs asssociated with recurring nonaudit services result in incremental economic bonding between an auditor and auditee. Such bonding, perceived as an impairment to auditor independence, could motivate an auditee to avoid such purchases, but the model predicts that cost savings from knowledge spillovers mitigate the level of bonding in a competitive market since auditors have an incentive to pass on some of the savings to the auditee. The auditor's limit-pricing scheme requires that the present value of the combined engagement fees be set marginally below the present value of both competitors' costs and the auditee's switching costs. The auditor thus offers a fee discount, which represents a portion of the cost savings from anticipated knowledge spillovers in a joint engagement.

Because the auditor offers a joint engagement discount, the auditee's decision involves a trade-off between the fee discount and the agency costs that could result from a perceived reduction in the quality of the independent audit.[5] In this multiperiod framework, Beck et al. (1988a) suggest that optimal joint production benefits and corresponding fee discounts are negotiated: auditors consider both start-up and normal costs for the expected levels of combined future services, and auditees consider anticipated agency costs when negotiating for the portfolio of audit and recurring nonaudit services. Because recurring services provide continuous rents to the auditor that can result in a perception of impaired independence, joint negotiations for audit and recurring nonaudit services are multiperiod decisions that affect the residual agency costs of joint production and limit the extent of the recurring services to be provided by the incumbent auditor.

---

[5] This view is consistent with that of DeAngelo (1981) and Beck et al. (1988a). Their analyses assume that auditees use a cost-benefit framework in making the decision to engage or switch auditors.

Further clarification of the auditee's decision framework involves a distinction between the manager-ownership firm and the external-ownership firm.[6] In the manager-ownership firm, an audit is demanded, if at all, for its internal control effects. The issue of independence does not arise, and recurring nonaudit services are procured for the benefits that accrue to the owner. If audit services are purchased, these benefits may include fee discounts that are offered as a result of knowledge spillovers from joint production. In the external-ownership firm, financial interests are diffused among shareholders and debtholders, and inherent agency conflicts influence the joint purchasing decision. In these firms, independent audits are demanded to attest to management's actions. Independence is now a consideration because the economic benefits of joint production are offset by agency costs that reduce the expected monitoring value of the audit. Thus, *ceteris paribus*, a lower level of nonaudit services would be demanded by these firms.

Nonrecurring Nonaudit Services

Nonrecurring nonaudit services generally are single engagements, but this category could also include related, sporadic engagements that occur on an irregular basis. The nature and extent of the knowledge spillovers from joint production and the related uncertainty level distinguish these services from recurring services. The model in Beck et al. (1988a) predicts that bonding effects of nonrecurring services would be minimal to nonexistent since future spillovers can be transitory and heavily discounted by future-period uncertainty. In contrast to the multiperiod framework for joint audit and recurring nonaudit services, a single-period decision suffices for a nonrecurring service. Fee negotiations occur as needed, separate from the continuing, jointly negotiated contracts.[7]

Hypotheses

When negotiating for both audit and recurring nonaudit services, auditees are expected to consider the agency costs associated with using the auditor in a long-term engagement. However, for nonrecurring services, auditees are expected to appraise the incremental agency costs as relatively small since the fee for a single-period service is minor in comparison with the present value of the recurring audit engagement. *Ceteris paribus*, it is hypothesized that auditees distinguish between recurring and nonrecurring services when they determine the level of nonaudit services to purchase from their auditor.

---

[6] In the manager-ownership firm, the manager owns 100 percent of the equity. In the external-ownership firm, financing is provided by external equity and debt (Jensen and Meckling 1976; Watts 1977).

[7] Several additional factors may favor an auditee's single-period decision to use the incumbent auditor when the need for nonrecurring nonaudit services arises. First, search costs for an alternate service provider can be avoided. Second, the incumbent auditor may be the low-cost or high-quality choice because of prior knowledge about the auditee or expertise within the auditee's industry. Although the model in Beck et al. (1988a) indicates that future knowlege spillovers are a necessary condition for fee discounts for nonrecurring nonaudit services, other factors could motivate the incumbent auditor to discount these services. For example, the search costs incurred for a nonaudit provider could reduce the auditee's future-period audit-switching costs and increase the likelihood of an auditor change. Thus, auditors would be motivated, even in the absence of future knowledge spillovers, to provide economic incentives to auditees and avoid impairing future-period audit negotiations.

Surrogates for agency costs include measures of management ownership, outside investment concentration, and debt (Chow 1982; Jensen and Meckling 1976; Niehaus 1989). Management ownership reduces agency conflict since a share of the equity is presumed to align the interests of management with those of the external owners. When managerial ownership is high, agency costs are likely to be lower, so auditees can take advantage of joint production benefits and purchase a higher level of nonaudit services. *Ceteris paribus*, it is hypothesized that the level of auditor-provided recurring nonaudit services is greater when managerial ownership is higher.

Diffused outside ownership provides little incentive for small individual owners to monitor management directly because limited benefits are received in comparison with the monitoring costs (Niehaus 1989). When outside investment is concentrated, however, agency conflict is reduced because a significant investment by an external owner is an incentive for direct monitoring that can supplement or substitute for the monitoring provided by the auditor. Hence, auditees can take advantage of joint production benefits and purchase higher levels of recurring nonaudit services from their auditor. *Ceteris paribus*, it is hypothesized that the level of auditor-provided recurring nonaudit services is greater when outside investment has higher concentration.

High levels of debt and accounting-based debt covenants are associated with the demand for audited financial statements (Chow 1982). Since debt can motivate wealth transfers through dividends or debt restructuring, the incentive to maintain independent monitoring should be greater when there is a higher level of debt. *Ceteris paribus*, it is hypothesized that the level of auditor-provided recurring nonaudit services is negatively related to leverage.

II. Method

ASR No. 250 (SEC 1978) required SEC registrants to disclose total nonaudit services provided by their external auditor as a percentage of audit fees and specific identification of services that were at least 3 percent of audit fees.[8] In this study, we examine data from *Fortune 500* companies for the entire three fiscal years (1978–80) during which ASR No. 250 was in effect. The sample consists of those auditees with data available from the following sources: the *Fortune 500 Corporate Data Exchange Stock Ownership Directory* for 1980 (CDE 1981), the 1985 version of COMPUSTAT, and proxy statements from a university library.

Ownership data for 1980 from the CDE excluded auditees that were acquired in 1980, those that were wholly or closely held by one or two companies, and those that were privately held. Ownership data for 1978 and 1979 were collected from proxy statements (the cutoff for external reporting of ownership was 5 percent of the outstanding

---

[8] Actual fee data for these companies are unavailable. However, survey research reports a positive relation between nonaudit and audit fees (Palmrose 1986b; Simunic 1984), and both Beck et al. (1988b) and Palmrose (1988) note that the ratio of nonaudit fees controls, in part, for company size differences.

Table 1
Description of the Variables

Dependent Variables:	
RECUR	Recurring nonaudit services (totals of *TAX*, *SYS*, and *PP*)
TAX	Tax service fees as a percentage of audit fees
PP	Pension and personnel fees as a percentage of audit fees
SYS	Systems fees as a percentage of audit fees
NONRECUR	Nonrecurring nonaudit services (totals of *MRGR* and *OTHER*)
MRGR	Merger and acquisition fees as a percentage of audit fees
OTHER	Other nonaudit fees as a percentage of audit fees
Independent Variables and Predicted Signs:	
MGRONR	Percentage of shares owned by managers and directors (Predictions: +with *RECUR*, not related to *NONRECUR*)
OUTCON	Market value of the percentage of shares owned by largest outside stockholder in log transformation (Predictions: +with *RECUR*, not related to *NONRECUR*)
LVRG	Ratio of book values of long-term debt to total assets (Predictions: −with *RECUR*, not related to *NONRECUR*)
Other Variables:	
NONADT	Total nonaudit services
AUDITOR	Takes value of 1 if the audit firm is Arthur Andersen and 0 otherwise
YEAR j	Two indicator variables for the years 1979 and 1980
INDSPC	Takes the value of 1 when the auditor is classified as an industry specialist and 0 otherwise
NETINC	Net income for current and two previous years deflated by total assets for the current year

shares). COMPUSTAT provided the financial data, and proxy statements identified the auditor and the level of purchased nonaudit services. Usable data from all these sources provided 860 observations for the three-year period.

Descriptions of the variables used are provided in table 1. To develop the dependent measures, disclosures of nonaudit services as a percentage of audit fees were initially divided into five categories: tax (*TAX*), pensions and personnel (*PP*), information systems (*SYS*), mergers and acquisitions (*MRGR*), and other nonaudit services (*OTHER*). In conformance with prior research (Beck et al. 1988b; DeBerg et al. 1991), recurring nonaudit services (*RECUR*) were measured by combining the totals for *TAX*, *PP*, and *SYS*.[9] Nonrecurring services (*NONRECUR*) combined the totals for both *MRGR* and *OTHER*.

---

[9] Definitions of recurring nonaudit services may differ. Most tax services obviously are recurring in nature, but other types of services may contain both recurring and nonrecurring components. To determine whether the classification scheme for recurring nonaudit services was appropriate for our data, we examined temporal correlation coefficients and obtained results comparable to those of Beck et al. (1988b). To further examine the classification scheme, we performed our multivariate regression analyses using various definitions of *RECUR*, which included the combinations of (1) *TAX* and *PP*, (2) *TAX* and *SYS*, (3) *TAX* and

The independent variables representing the agency measures included managerial ownership, outside investment concentration, and leverage. Managerial ownership (*MGRONR*) was the percentage of outstanding shares owned by managers and directors. Outside investment concentration (*OUTCON*) was (a log transformation of) the market value of outstanding shares owned by the largest outside shareholder. The proxy for leverage (*LVRG*) was the ratio of book value of long-term debt to book value of total assets.[10]

Prior research (Beck et al. 1988b; Palmrose 1986a; Scheiner and Kiger 198[illegible]) on nonaudit services suggests several factors that could affect purchase. Examination of these factors led to the inclusion of several control measures, including variables to represent auditor (*AUDITOR*) and industry specialist classifications (*INDSPC*). In addition, two indicator variables were used for 1979 and 1980 observations to denote changes in the demand for nonaudit services during this period. Finally, a control variable measuring net income (*NETINC*) was included to proxy for the availability of resources that could affect the purchase of nonaudit services.

Regression analysis was used to examine the hypothesized relation with *MGRONR*, *OUTCON*, and *LVRG* after controlling for various factors that could affect the purchase of nonaudit services. To estimate parameters, we combined time-series and cross-sectional data for the sample auditees and adopted a model with constant slope coefficients and varying intercepts over different years (Judge et al. 1980, 338).[11] The following form of regression model was estimated to test the stated hypotheses:

$$\begin{aligned}\text{Nonaudit Services} = {} & b_0 + b_1 MGRONR + b_2 OUTCON + b_3 LVRG \\ & + b_4 AUDITOR + b_5 YEAR1979 + b_6 YEAR1980 \\ & + b_7 INDSPC + b_8 NETINC + e. \qquad (1)\end{aligned}$$

Two different regressions using equation (1) were estimated with nonaudit services alternately taking the values of *RECUR* and *NONRECUR*. With *RECUR* as the dependent variable, the coefficients b_1, and b_2 were predicted to be positive, and b_3 was predicted to be negative. With *NONRECUR* as the dependent variable, b_1, b_2, and b_3 were all predicted to be equal to zero. To test the stability of estimated regression coefficients, we also divided our sample into two subsamples based on size, as measured by the auditee's total assets, then estimated two additional regressions with *RECUR* as the

---

MRGR, and (4) *TAX*, *PP*, *SYS*, and *MRGR*. These results (not reported) were similar to those presented. From this sensitivity analysis, we infer that the results for recurring nonaudit services are robust to slight deviations in the definition of *RECUR*. For comparability with earlier research (Beck et al. 1988b; DeBerg et al. 1991) we report the results for *RECUR* using the originally specified definition.

[10] When these variables were defined in several other ways, the results (not reported) were qualitatively unchanged. For example, using the CDE data set for 1980, we respecified *MGRONR* as the number of managers and directors among the top five stockholders, and *OUTCON* as the percentage of outstanding shares owned by outsiders among the top five stockholders. *MGRONR* also was defined as managers only (excluding directors), *OUTCON* was defined as the percentage owned by the largest shareholder, and *LVRG* was defined as the ratio of the book value of long-term debt to the book value of equity.

[11] We also considered a regression model to incorporate slope shifts in the different years. For that, we included terms to measure the interaction of the dummy variables for 1979 and 1980 with *MGRONR*, *OUTCON*, and *LVRG*. Estimated coefficients (not reported) of all six interaction terms were statistically insignificant. From these results, we infer an absence of structural shifts in the various years and present the empirical analysis without interaction terms.

dependent variable.[12] Subsample 1 included the observations below the median size, and subsample 2 contained those above the median.

III. Results

Panel A of table 2 presents the distribution of total, recurring, and nonrecurring nonaudit services disclosures by auditor. The data indicate that Big Eight auditors were used by more than 95 percent of the sample, with the remaining observations also including national firms (i.e., those ranked within the next five largest audit firms). Among the Big Eight observations, Arthur Andersen auditees report the highest means for both total (43.92 percent) and recurring (37.34 percent) nonaudit services. For those with other Big Eight auditors, *NONADT* ranges from 17.04 to 23.99 percent and *RECUR* ranges from 12.21 to 15.73 percent.[13] The distinctive characteristics of nonaudit services procured by clients of Arthur Andersen were considered by including an indicator variable (*AUDITOR*) in the multivariate analyses. In a later section, we examine the effect of excluding Arthur Andersen auditees from the sample.[14]

Panel B in table 2 provides the distribution of nonaudit services by year (1978–80). In conformance with the SEC staff study reported in ASR No. 296 (SEC 1981) and Beck et al. (1988b), acquisitions of *NONADT* and *RECUR* consistently decline over the three-year period. From 1978 to 1979, the mean value of *NONADT* (*RECUR*) decreases from 26.99 to 22.88 (20.60 to 17.58) percent. In 1980, the value again declines to 21.98 (16.04) percent.[15] For *NONRECUR*, however, the results are different. Although the 1978 mean value of 6.38 percent initially declines in 1979 to 5.30 percent, it then increases to 5.94 percent in 1980. To control for the change in nonaudit services, two indicator variables were included in the multivariate analysis to represent 1979 and 1980 observations.

Table 3 reports the distribution of recurring nonaudit services by auditor and industry membership for seven major nonregulated industry classifications based on two-digit SIC codes. Although there is no difference between industries, there is a distinctive pattern within each industry classification.[16] Consistent with prior audit fee

---

[12] We also estimated regression (1) with *RECUR* as the dependent variable and auditee size as an additional independent variable. According to the results (not reported), the coefficient of size was insignificant, and the significance of *OUTCON* was reduced marginally compared to the results reported in table 6. Multicollinearity was high, as indicated by the condition index of 9.8 (Belsley et al. 1980). Therefore, we decided to report results without firm size.

[13] We explored whether there were significant differences in *NONADT* or *RECUR* for the various auditors. Two Kruskal-Wallis one-way ANOVA tests were used to determine if the nine subsamples of auditors (Big Eight and other) had the same distributions of *NONADT* and *RECUR*. The null hypotheses for same distributions were rejected on the basis of chi-square statistics of 76.31 (*NONADT*) and 81.25 (*RECUR*), which were significant at better than 0.01 levels. The ANOVA tests were repeated for seven subsamples (excluding Arthur Andersen and the others category), and the null hypotheses for same distributions of *NONADT* and *RECUR* could not be rejected. These results were further confirmed by two Kolmogorov-Smirnov sample tests comparing Arthur Andersen with all other auditors. The statistics of 3.592 and 3.764 ($p < 0.01$, two-tailed tests) reject the null hypotheses that Arthur Andersen and non-Arthur Andersen auditees have the same distributions of both total and recurring nonaudit services.

[14] Our tests were repeated after excluding the 35 observations with non-Big Eight auditors. The regression results were similar, and the adjusted R-square improved only marginally.

[15] Two Kruskal-Wallis one-way ANOVA tests indicate that the three subsamples do not have the same distributions of *NONADT* and *RECUR*. The chi-square statistics of 15.72 and 13.52 are both significant at better than 0.01 levels.

[16] We could not reject the null hypothesis of no difference in recurring nonaudit distributions across industry groupings on the basis of a Kruskal-Wallis one-way ANOVA. The average recurring nonaudit services were in a narrow range across different nonregulated industry groups (16.51 to 20.78 percent).

Table 2
1978–80 Distribution of Nonaudit Services by Auditor and Year

			Total Nonaudit Fees as a Percentage of Audit Fees					*Recurring Nonaudit Fees as a Percentage of Audit Fees*					*Nonrecurring Nonaudit Fees as a Percentage of Audit Fees*				
	N	% of Total	*Mean*	*S.D.*	*Median*	*Min.*	*Max.*	*Mean*	*S.D.*	*Median*	*Min.*	*Max.*	*Mean*	*S.D.*	*Median*	*Min.*	*Max.*
Panel A. Distribution by Auditor:																	
Arthur Andersen	143	16.6	43.92	43.87	30.0	2	268	37.34	40.95	24.3	0	263	6.66	3.67	3.7	0	73
Arthur Young	76	8.9	19.05	13.21	17.0	0	52	14.60	11.90	12.0	0	49	4.45	6.77	2.0	0	41
Coopers & Lybrand	84	9.8	23.99	19.88	17.5	2	100	15.73	13.43	12.0	0	52	8.26	13.43	4.0	0	70
Deloitte Haskins & Sells	88	10.2	20.60	18.94	15.5	0	150	14.63	15.81	11.0	0	115	5.97	9.63	2.8	0	66
Ernst & Whinney	122	14.2	17.04	12.27	14.0	2	57	12.21	10.58	10.0	0	56	4.83	6.81	3.0	0	43
Peat Marwick	107	12.4	19.20	12.28	17.0	0	62	15.07	12.38	13.7	0	62	4.13	5.71	2.0	0	34
Price Waterhouse	167	19.4	20.81	17.19	16.0	1	108	15.02	15.57	12.0	0	102	5.80	8.50	3.0	0	48
Touche Ross	38	4.4	23.13	31.80	16.0	2	193	12.74	8.78	11.0	2	32	10.40	29.55	2.0	0	176
Others	35	4.1	11.29	9.98	8.0	2	53	7.52	6.87	5.3	0	24	3.77	6.79	2.0	0	38
Totals	860	100.0	23.78	25.29	17.0	0	268	17.93	22.30	12.0	0	263	5.86	10.51	3.0	0	176
Panel B. Distribution by Year:																	
1978	257	29.88	26.99	26.96	20.0	0	195	20.60	23.28	15.0	0	195	6.38	13.91	2.7	0	176
1979	291	33.84	22.88	27.18	15.0	0	268	17.58	25.09	11.0	0	263	5.30	8.90	2.5	0	66
1980	312	36.28	21.98	21.62	17.0	0	207	16.04	18.19	11.0	0	134	5.94	8.46	3.0	0	73

Table 3

1978–80 Distribution of Recurring Nonaudit Services by Industry and Auditor

	Industries							
	*IND*1	*IND*2	*IND*3	*IND*4	*IND*5	*IND*6	*IND*7	*Auditor Means*
Arthur Andersen	61.73*	40.92*	31.67*	28.56*	22.00	36 .04*	70.08*	37.34
Arthur Young	12.97	7.30	18.38*	15.71	32.00*	12.17	5.75	14.60
Coopers & Lybrand	7.67	11.81	19.21*	15.87	27.44*	10.42	8.60	15.73
Deloitte Haskins & Sells	15.35	25.00*	10.68	11.38	19.67	13.72	31.67*	14.63
Ernst & Whinney	8.64	15.20	14.00	9.21	8.13	17.58*	2.25	12.21
Peat Marwick	7.75	7.10	0.67	18.15*	20.79	14.52	13.75	15.07
Price Waterhouse	17.13*	11.93	9.89	13.75	29.63*	12.11	14.02	15.02
Touche Ross	14.83	14.69	15.67	0.00	13.33	11.30	8.50	12.74
Others	7.67	7.33	0.00	8.59	2.83	9.14	2.94	7.52
Industry Means	17.67	18.77	17.35	16.51	20.78	17.38	20.22	17.93

Note: An * denotes use of an industry specialist as auditor. The industry classifications are: *IND*1, Metal, Ores, Mining Works and Products; *IND*2, Consumer Goods (Food, Drinks, Tobacco, Textile, Apparel, etc.); *IND*3, Paper and Publishing; *IND*4, Petroleum and Chemicals; *IND*5, Electric Machinery; *IND*6, Motor Vehicles and Other Machinery; and *IND*7, Other Nonregulated.

research that classified auditors as industry specialists to reflect differential capacity and quality of audit services (Palmrose 1986a), our classification of industry specialists (*INDSPC*) was based on industry participation within the available sample. *INDSPC* includes the auditor with the largest mean recurring nonaudit services measure within each industry. Also included are the second and possibly third largest providers when distinctive differences were observed between the second and third or between the third and remaining providers. These classification rules identified two industry specialists within five of the industries and three within the remaining two industries. *INDSPC* takes the value of 1 when the auditor is an industry specialist, zero otherwise.

Descriptive Statistics

Table 4 contains descriptive data on the total disclosures of procured nonaudit services and the various subcategories of these services. Although more than 98 percent of the sample procured some type of nonaudit service from their auditor, all subclassifica-

Table 4
Percentage of Companies Reporting Nonaudit Services and Descriptive Statistics

Panel A. Percentage of Companies Reporting Nonaudit Services:

Type of Fee	Range of the Reported Fees							
	None	<3%	3%–9.99%	10%–24.99%	25%–49.99%	50%–74.99%	75+%	Total Sample
RECUR	6.16	7.56	25.58	39.19	15.81	2.91	2.79	100%
	(53)	(65)	(220)	(337)	(136)	(25)	(24)	(860)
TAX	10.00	9.77	35.70	33.83	8.72	1.05	0.93	100%
	(86)	(84)	(307)	(291)	(75)	(9)	(8)	(860)
PP	70.70	5.81	19.07	4.30	0.12	0	0	100%
	(608)	(50)	(164)	(37)	(1)	(0)	(0)	(860)
SYS	65.12	7.33	16.27	6.74	2.79	0.70	1.05	100%
	(560)	(63)	(140)	(58)	(24)	(6)	(9)	(860)
NONRECUR	26.05	21.86	34.53	13.37	3.49	0.58	0.12	100%
	(224)	(188)	(297)	(115)	(30)	(5)	(1)	(860)
MRGR	75.35	5.93	13.25	3.84	1.63	0	0	100%
	(648)	(51)	(114)	(33)	(14)	(0)	(0)	(860)
OTHER	34.42	24.42	30.93	8.25	1.28	0.58	0.12	100%
	(296)	(210)	(266)	(71)	(11)	(5)	(1)	(860)
NONADT	1.05	3.37	20.23	42.67	22.79	6.40	3.49	100%
	(9)	(29)	(174)	(367)	(196)	(55)	(30)	(860)

Panel B. Descriptive Statistics:

	Mean	Std. Dev.	Median	Min.	Max.
RECUR	17.93	22.30	12	0	263
TAX	11.69	14.45	8	0	162
PP	1.68	3.65	0	0	40
SYS	4.56	14.79	0	0	177
NONRECUR	5.86	10.51	3	0	176
MRGR	1.85	5.37	0	0	46
OTHER	4.01	8.67	2	0	155
NONADT	23.78	25.29	17	0	268

Note: All variables are defined in table 1.

tions of *NONADT* have highly skewed distributions. *RECUR* was purchased by more than 93 percent of the sample, with its subcategories of *TAX*, *PP*, and *SYS* being reported by 90, 29, and 34 percent, respectively. *NONRECUR* was procured by approximately 74 percent of the sample, with its subcategories of *MRGR* and *OTHER* being reported by 24 and 65 percent. In particular, the frequency distribution for *TAX*

Table 5
Descriptive Statistics and Correlation Coefficients Between Variables

Panel A. Descriptive Statistics, Independent Variables:

	Mean	*Std. Dev.*	*Median*	*Min.*	*Max.*
MGRONR	7.21	12.69	0.00	0.00	64.99
OUTCON	3.70	1.30	3.70	0.08	9.35
LVRG	0.18	0.10	0.18	0.00	0.56

Panel B. Pearson Correlation Coefficients, Independent Variables:

	MGRONR	*OUTCON*	*LVRG*
MGRONR	—		
OUTCON	−0.1812*	—	
LVRG	−0.1436*	−0.2089*	—

Panel C. Pearson Correlation Coefficients, Independent and Dependent Variables:

		Dependent Variables for Recurring Services				*Dependent Variables for Nonrecurring Services*		
	Predicted	*RECUR*	*TAX*	*PP*	*SYS*	*NON-RECUR*	*MRGR*	*OTHER*
MGRONR	+	0.0548+	0.0613+	0.0156	0.0186	0.0967+	0.1210+	0.0423
OUTCON	+	0.1007+	0.1208+	0.0389	0.0499+	−0.0413	0.0038	−0.0505
LVRG	−	−0.0992+	−0.1199+	−0.0438+	−0.0512+	−0.0073	−0.0236	−0.0057

Note: All variables are defined in table 1.

* Significant at $p<0.10$, two-tailed test.
+ Significant at $p<0.10$, one-tailed test.

is much higher than the frequencies for any other specific nonaudit service. While *TAX* has 92 observations with values of 25 percent or higher, there is only one for *PP*, 17 for *OTHER*, 14 for *MRGR*, and 39 for *SYS*. Although it is not possible to compare our sample directly with that of Beck et al. (1988b), there are similar patterns in the data that support their suggestion of varying demand for nonaudit services.

Descriptive statistics and Pearson correlation coefficients for the independent variables are reported in panels A and B of table 5. *MGRONR* and *OUTCON* are negatively correlated, which indicates that high managerial ownership is associated with a lower value of holdings by the largest outside shareholder. The largest correlation is between *OUTCON* and *LVRG*, but the magnitude of the correlation coefficients of various independent variables with one another is rather low (−0.1436 to −0.2089). Therefore, multicollinearity is unlikely to be a problem in interpreting the results of multivariate analysis. We further examine multicollinearity in the discussion of the regression results.

Univariate Analysis

Panel C of table 5 contains the correlation coefficients between the nonaudit services measures and the independent variables. The signs of the correlation coefficients of *RECUR* with *MGRONR*, *OUTCON*, and *LVRG* are as predicted and are significant at 0.10 levels (one-tailed tests). For *NONRECUR*, the correlation coefficients are negative with *OUTCON* and *LVRG*, but both coefficients are insignificant, as predicted. *MGRONR* has a positive and significant correlation coefficient with *NONRECUR*, whereas it was predicted to have no significant relationship. To further examine these results, the correlation coefficients of the independent variables with the individual nonaudit services components also are reported. The signs of the correlation coefficients of the recurring categories of *TAX*, *PP*, and *SYS* with the independent variables are consistent with the composite measure of *RECUR*. The unexpected significant relation between *NONRECUR* and *MGRONR* is duplicated for the nonrecurring category of *MRGR*, but not for the *OTHER* category.

Multivariate Analysis

The results of the estimation of regression (1) with *RECUR* and *NONRECUR* as the dependent variables are reported in table 6. Also included are the results for the two subsamples based on size with *RECUR* as the dependent variable. White's (1980) test rejected the null hypothesis that the regression residuals are homoskedastic. Therefore, we report adjusted *t*-statistics based on the heteroskedasticity-consistent covariance matrix (White 1980). Hypothesis testing also assumes a normal distribution, but the residuals and the dependent variable (and its transformations not reported here) are not normally distributed. Hence, *t*-statistics based on the jackknife distribution-free nonparametric estimate of standard error are also included (Miller 1974a, 1974b). We also assessed the effect of multicollinearity by applying the procedures suggested by Belsley et al. (1980); for the results in table 6, the highest condition index was 2.4, which indicates that the regressors do not involve any significant dependencies.

The estimated model with *RECUR* as the dependent variable has an adjusted R^2 of 0.2604 and an *F*-value of 36.623, significant at a 0.0001 level.

The results for *MGRONR*, *OUTCON*, and *LVRG* are consistent with the univariate results in panel C of table 5, and their coefficients have the predicted signs at significance levels of 0.01 (one-tailed tests) for *OUTCON* and *LVRG* and 0.05 (one-tailed test) for *MGRONR*. The coefficients of *INDSPC* and *NETINC* are positive and significant at better than a 0.05 level (two-tailed test). These results imply that firms with industry specialists procure more recurring nonaudit services from their auditors than firms not using the services of industry specialists. These results also provide empirical evidence that the greater the resources, as represented by net income, the higher the recurring nonaudit services procured from the auditor.

The *AUDITOR* coefficient is positive and significant at a 0.05 level (two-tailed test), which indicates that Arthur Andersen clients procure more recurring nonaudit services than clients of all other audit firms.[17] In the first round of outlier isolation, seven of ten outliers were clients of Arthur Andersen, as were nine of the next ten isolated outliers. These particular observations as well as the total 20 identified outliers are

---

[17] A partial F-test indicates that the *AUDITOR* variable adds significant explanatory power (0.01 level).

Table 6
Regression Results of Recurring and Nonrecurring Nonaudit Services

	Intercept	*MGRONR*	*OUTCON*	*LVRG*	*AUDITOR*	*YEAR1979*	*YEAR1980*	*INDSPC*	*NETINC*	*# of Obser-vations*	*Adjusted R-Square*	*F-Value (Signif-icance)*
Dependent Variable: *RECUR* (Predictions)		+	+	–	+	–	–	+	+			
Coefficient	5.171	0.881	2.153	−19.372	19.265	−3.102	−4.185	8.408	2.47	860	0.2604	36.623
Adjusted *t*-value	1.52	2.20*	4.48**	−2.56**	6.69+	−2.21+	−3.79+	4.72+	5.21+			(0.0001)
Jackknife *t*-value	1.49	2.17*	4.43**	−2.54**	6.60+	−2.21+	−2.75+	4.65+	5.18+			
Dependent Variable: *NONRECUR* (No predicted relation)												
Coefficient	6.606	0.078	−0.145	−0.691	1.931	−1.081	−0.484	−1.508	0.116	860	0.0068	1.737
Adjusted *t*-value	3.47+	2.36**	−0.43	−0.15	1.42	−1.09	−0.58	−1.46	0.52			(0.0863)
Jackknife *t*-value	3.43+	2.12*	−0.42	−0.09	1.43	−1.06	−0.58	−1.40	0.50			
Dependent Variable: *RECUR* (Predictions)		+	+	–	+	–	–	+	+			
Subsample 1 (Total assets less than median)												
Coefficient	3.723	0.184	2.161	−17.634	21.075	−2.430	−4.823	6.919	2.707	430	0.2906	21.864
Adjusted *t*-value	0.72	2.47**	2.09*	−2.46**	5.18+	−1.17	−2.12+	2.78+	6.41+			(0.0001)
Jackknife *t*-value	0.70	2.43**	2.06*	−2.39**	4.78+	−1.08	−2.23+	2.73+	6.37+			
Subsample 2 (Total assets greater than median)												
Coefficient	10.263	0.079	2.266	−24.053	16.668	−3.593	3.898	9.426	1.861	430	0.2249	16.246
Adjusted *t*-value	1.98+	1.89*	3.16**	−2.74**	4.56+	−2.32+	−1.71	3.85+	2.57+			(0.0001)
Jackknife *t*-value	1.90	1.85*	3.14**	−2.70**	4.42+	−2.40+	−1.75	3.82+	2.49+			

Note: All variables are defined in table 1. Adjusted *t*-values are based on White (1980).

* Significant at $p<0.05$, one-tailed test.
+ Significant at $p<0.05$, two-tailed test.
** Significant at $p<0.01$, one-tailed test.

evenly distributed over different industries and years. The values of *RECUR* range from 79 to 263 percent for the 20 outliers. Keeping in mind the distinctive characteristics of the Arthur Andersen clients, we tested whether these had different distributions of *MGRONR*, *OUTCON*, and *LVRG* from clients of other auditors. The mean values of *MGRONR* were 7.53 for Authur Andersen clients and 7.14 for others. *OUTCON* mean values were 3.69 for Arthur Andersen clients and 3.70 for others. The *LVRG* means were 0.174 (Arthur Andersen clients) and 0.181 (others). Kolmogorov-Smirnov statistics of 1.45 and 1.38 ($p<0.05$, two-tailed test) for *MGRONR* and *LVRG*, respectively, indicated that we could reject the null hypotheses that clients of Arthur Andersen and clients of all other audit firms have the same distributions of *MGRONR* and *LVRG*. The Kolmogorov-Smirnov statistic for *OUTCON* is not significant at any conventional level. Therefore, it could be concluded that, on average, Arthur Andersen clients have significantly higher *MGRONR* and significantly lower *LVRG*.

The results of the estimated model with *NONRECUR* as the dependent variable (table 6) are again consistent with earlier analysis. The adjusted R^2 and *F*-values for *NONRECUR* are 0.0068 and 1.737. As predicted, all variables are insignificant for *NONRECUR* as the dependent variable with the exception of the *MGRONR* coefficient, which is positive and significant.

Table 6 also includes regression results for the two subsamples based on auditee's size, to assess the homogeneity of regression models. The results are qualitatively the same as for the total sample with *RECUR* as the dependent variable. However, the 1979 year indicator variable is significant only for the larger auditees, and the 1980 year indicator variable is significant only for the smaller auditees. The computed *F*-statistic (1.06) to test differential regressions in the subsamples is insignificant at any conventional level, and we cannot reject the null hypothesis of homogeneous regressions in the two groups.

The results of the estimation of regression (1) with the individual components of *RECUR* and *NONRECUR* as the dependent variables are reported in table 7. The results pertaining to *OUTCON* and *LVRG* are consistent with earlier analysis. For the independent variable *MGRONR*, the results are mixed. The coefficient *MGRONR* is positive, as expected, when the dependent variable is *TAX*, *PP*, or *SYS*, but it is significant only when *TAX* is the dependent variable. When *MRGR* is the dependent variable, *MGRONR* has a positive and significant coefficient (0.01 level, one-tailed test) which is inconsistent with the prediction, However, with *OTHER* as the dependent variable, the coefficient of *MRGR* is insignificant, as expected. Hence, the procurement of nonaudit merger and acquisition services is related to a higher level of management ownership. These results combined with those of table 6 suggest that managerial ownership has an effect on nonrecurring nonaudit services only if *MRGR* is included.

Further Analyses

To test the robustness of the results with the exclusion of Arthur Andersen auditees and to adjust for serial correlation in error terms because of pooling three years data, we selected non-Arthur Andersen auditees with data available in all three years. We estimated regression equation (1) excluding the variable *AUDITOR*. The results shown in panel A of table 8 are qualitatively the same as those in table 6 with *RECUR* as the dependent varaible. The model has an adjusted R^2 of 0.1389 and an *F*-value of 15.42, significant at a 0.0001 level.

Table 7
Regression Results of Specific Categories of Nonaudit Services

	Intercept	*MGRONR*	*OUTCON*	*LVRG*	*AUDITOR*	*YEAR1979*	*YEAR1980*	*INDSPC*	*NETINC*	*# of Obser-vations*	*Adjusted* R-*Square*	*F-Value (Signif-icance)*
Recurring Services (Predictions)		+	+	−	+	−	−	+	+			
Tax												
Coefficient	3.624	0.701	1.103	−10.685	8.390	0.795	1.231	6.035	0.405	860	0.1591	20.163
Adjusted *t*-value	1.53	2.28*	4.88**	−2.64**	3.93+	0.54	0.86	4.86+	1.39			(0.0001)
Jackknife *t*-value	1.51	2.19*	4.83**	−2.59**	3.87+	0.50	0.87	4.79+	1.42			
Pension/Personnel												
Coefficient	4.183	0.139	0.491	−2.119	0.708	−3.679	−3.862	0.318	0.132	860	0.2578	36.634
Adjusted *t*-value	6.82+	1.56	2.13*	−1.86*	0.22	−10.42+	−11.06+	0.96	1.96+			(0.0001)
Jackknife *t*-value	6.81+	1.59	2.01*	−1.92*	0.18	−9.31+	−10.97+	0.82	1.89+			
Systems												
Coefficient	−2.635	0.041	0.559	−6.567	10.166	−0.218	−1.554	2.055	1.939	860	0.1656	22.431
Adjusted *t*-value	−1.01	1.01	2.55**	−2.37**	4.87+	−0.08	−1.08	2.09+	7.03+			(0.0001)
Jackknife *t*-value	−1.00	0.97	2.49**	−2.36**	4.42+	−0.12	−0.96	2.04+	6.58+			
Nonrecurring Services (No predicted relation)												
Mergers and Acquisitions												
Coefficient	1.874	0.054	0.136	−0.528	0.737	−0.772	−1.244	−0.883	0.085	860	0.0209	3.568
Adjusted *t*-value	2.08+	3.29**	0.83	−0.19	0.96	−1.52	−2.32+	−1.83	0.69			(0.0002)
Jackknife *t*-value	1.96+	3.11**	0.84	−0.21	0.88	−1.51	−2.29+	−1.80	0.76			
Other Services												
Coefficient	4.732	0.024	−0.281	−0.164	1.194	−0.309	0.760	−0.624	0.031	860	−0.0008	0.917
Adjusted *t*-value	3.03+	0.64	−1.08	−0.04	1.31	−0.47	1.06	−0.79	0.17			(0.5015)
Jackknife *t*-value	3.00+	0.58	−1.06	−0.05	1.30	−0.49	1.08	−0.78	0.15			

Note: All variables are defined in table 1. Adjusted *t*-values are based on White (1980).

* Significant at $p<0.05$, one-tailed test.
+ Significant at $p<0.05$, two-tailed test.
** Significant at $p<0.01$, one-tailed test.

Table 8

Regression Results of Recurring Nonaudit Services for Non-Arthur Andersen Auditees

Panel A. Dependent Variable RECUR (Predictions):

	Intercept	*MGRONR*	*OUTCON*	*LVRG*	YEAR1979	YEAR1980	INDSPC	NETINC	*Number of* Observations	*Adjusted* R-Square	F-*Value* (*Significance*)
		+	+	−	−	−	+	+			
Coefficient	10.487	0.098	1.633	−14.611	−3.406	−4.116	0.981	2.147	633	0.1389	15.42
Adjusted *t*-value	4.39+	2.23*	3.58**	−2.61**	−2.73+	−3.28+	5.06+	4.18+			(0.0001)
Jackknife *t*-value	4.77+	2.20*	3.48**	−2.84**	−2.82+	−3.54+	4.96+	3.54+			

Panel B. Transformed Dependent Variable TRECUR (Predictions):

	Intercept	*TMGRONR*	*TOUTCON*	*TLVRG*	*INDSPC*	*TNETINC*	*Number of Observations*	*Adjusted* R-*Square*	F-*Value* (*Significance*)
		+	+	−	+	+			
Coefficient	2.661	0.119	2.416	−13.430	6.532	1.867	633	0.1131	17.12
Adjusted *t*-value	1.40	2.09*	4.91**	−1.97*	4.62+	3.01+			(0.0001)
Jackknife *t*-value	1.49	2.06*	4.77**	−2.08*	4.51+	2.98+			

Note: The variables are defined in table 1. The T-prefix indicates transformed variables; see equations (2) and (3). Adjusted *t*-values are based on White (1980).

* Significant at $p<0.05$, one-tailed test.
+ Significant at $p<0.05$, two-tailed test.
** Significant at $p<0.01$, one-tailed test.

The application of ordinary least squares to estimate regression equation (1) with serial correlation in error terms would give unbiased but inefficient estimates.[18] It has been suggested (Johnston 1984, chap. 8) that unbiased and efficient estimation can be achieved by using a generalized least squares transformation, which eliminates the above problem through the following transformation:

$$\begin{aligned} TVAR_t &= \sqrt{1-\rho^2}VAR_t, && \text{for } t=1, \text{ and} \\ &= VAR_t - \rho VAR_{t-1}, && \text{for } t=2 \text{ and } 3, \end{aligned} \tag{2}$$

where ρ is first-order serial correlation coefficient, *VAR* represents the variables *RECUR*, *MGRONR*, *OUTCON*, *LVRG*, and *NETINC*. Therefore, we estimated the following regression equation:

$$\begin{aligned} TRECUR = {} & c_0 + c_1 TMGRONR + c_2 TOUTCON + c_3 TLVRG + c_4 INDSPC \\ & + c_5 TNETINC + \epsilon, \end{aligned} \tag{3}$$

where *TRECUR*, *TMGRONR*, *TOUTCON*, *TLVRG*, and *TNETINC* are the transformed values of *RECUR*, *MGRONR*, *OUTCON*, *LVRG*, and *NETINC*, respectively, from equation (2), and ϵ is the serially independent error term.

The results of the estimation of regression (3) are reported in panel B of table 8 and are qualitatively similar to those achieved with the original variables (panel A, table 8). As expected, the adjusted *R*-square is reduced, from 0.1389 to 0.1131, after the transformation.

IV. Summary and Concluding Remarks

Overall, the results of this study suggest that auditees recognize the potential for perceptions of independence impairment and voluntarily manage the amount of recurring nonaudit services that are purchased from their auditor. These results also support the implications of knowledge spillover from the Beck et al. (1988a) model of the market for joint audit and nonaudit services. Additionally, evidence that auditees manage joint engagements in accordance with potential agency costs is consistent with and explains Glezen and Millar's (1985) earlier empirical conclusion that the stockholders in their study did not perceive that auditor independence was affected by the reported levels of nonaudit services.

A limitation of this study is that operationalization of recurring and nonrecurring services are difficult and may require finer partitioning. When five subclassifications of nonaudit services are examined separately, the effects of managerial ownership are mixed. Managerial ownership unexpectedly is asssociated with nonrecurring merger and acquisition services and is not associated with recurring systems or pension and personnel services. One explanation may be that these services have both recurring and nonrecurring components, as well as other possible multiperiod outcomes (e.g., for mergers and acquisitions, owner-managers may expect significant, long-term employment effects). Similarly, tax services can include both recurring and nonrecurring elements. If certain tax compliance work is so closely related to an audit as to be virtually inseparable for large companies, an alternative explanation for the tax category effects could be that the coefficients represent sources of underlying demand for certain tax services.

---

[18] The sampling variances are biased and likely to be understated, thus, an exaggerated R^2 results (see Maddala 1988, chap. 6).

Another limitation is that this study was restricted to publicly available data, and no inferences can be made regarding agency costs and nonaudit services subsequent to the mandatory disclosure period or beyond the bound of *Fortune 500* companies. However, in a survey of audit committee chairmen from *Fortune 1000* industrial and service companies (POB 1986), 20 percent of the respondents reported taking the position that their auditor should not be used for a proposed nonaudit assignment because it could impair (or be perceived as impairing) independence. Thus, within this group, audit committees may be an important mechanism to monitor and limit purchases of nonaudit services.

Another finding from this study is that auditees who used the services of industry specialists or Arthur Andersen exhibited significantly different procurement patterns and a stronger demand for recurring nonaudit services from their auditor. One explanation for this may be that auditor selection explicitly includes an evaluation of an auditor's ability to provide a particular type or level of recurring nonaudit services. Another possible interpretation is that auditees perceive the acceptable level of recurring nonaudit services according to their chosen auditor. Since changes in the need for nonaudit services or a reorganization of ownership and capital structure can alter current joint services contracts and the economic benefits of joint production, auditees may be motivated to switch auditors for nonaudit services. These potential explanations and rapid growth in the market for nonaudit services suggest that future research should examine the significance of nonaudit services as a factor in auditor-switching behavior.

References

Abdel-khalik, A. R. 1990. The jointness of audit fees and demand for MAS: A self-selection analysis. *Contemporary Accounting Research* 6 (Spring): 295–322.

Antle, R. 1984. Auditor independence. *Journal of Accounting Research* 22 (Spring): 1–20.

Beck, P. J., T. J. Frecka, and I. Solomon. 1988a. A model of the market for MAS and audit services: Knowledge spillovers and auditor-auditee bonding. *Journal of Accounting Literature* 7: 50–64.

——, ——, and ——. 1988b. An empirical analysis of the relationship between MAS involvement and auditor tenure: Implications for auditor independence. *Journal of Accounting Literature* 7: 65–84.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Data and Sources of Collinearity.* New York: Wiley.

Berton, L. 1989. SEC is studying independence of firms doing consulting and accounting work. *The Wall Street Journal,* November 10, A12.

Chow, C. W. 1982. The demand for external auditing: Size, debt and ownership influences. *The Accounting Review* 57 (April): 301–26.

Commission on Auditors' Responsibilities. 1978. *The Commission on Auditors' Responsibilities: Report, Conclusions, and Recommendations.* New York: CAR.

Corporate Data Exchange, Inc. 1981. *CDE Stock Ownership Directory, No. 5, Fortune 500.* New York: CDE.

DeAngelo, L. 1981. Auditor independence, low-balling, and disclosure regulation. *Journal of Accounting and Economics* 3 (August): 113–27.

DeBerg, C. L., S. E. Kaplan, and K. Pany. 1991. An examination of some relationships between nonaudit services and auditor change. *Accounting Horizons* 5 (March): 17–28.

Fama, E. F., and M. C. Jensen. 1983a. Separation of ownership and control. *Journal of Law and Economics* 26 (June): 301–26.

——, and ——. 1983b. Agency problems and residual claims. *Journal of Law and Economics* 26 (June): 327–50.

Glezen, G. W., and J. A. Millar. 1985. An empirical investigation of stockholder reaction to disclosures required by ASR No. 250. *Journal of Accounting Research* 23 (Autumn): 859–70.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (October): 305–60.

Johnston, J. 1984. *Econometrics Methods.* New York: McGraw-Hill.

Judge, G., R. Hill, W. Griffiths, and T. Lee. 1980. *The Theory and Practice of Econometrics.* New York: Wiley.

Maddala, G. S. 1988. *Introduction to Econometrics.* New York: Macmillan.

Mednick, R. 1990. Independence: Let's get back to basics. *Journal of Accountancy* 169 (January): 86–93.

Miller, R. G. 1974a. The jackknife—A review. *Biometrika* 61 (April): 1–16.

———. 1974b. An unbalanced jackknife. *Annals of Statistics* 2 (September): 880–91.

Niehaus, G. 1989. Ownership structure and inventory method choice. *The Accounting Review* 64 (April): 269–84.

Palmrose, Z. V. 1986a. Audit fees and auditor size: Further evidence. *Journal of Accounting Research* 24 (Spring): 97–110.

———. 1986b. The effect of nonaudit services on the pricing of audit services: Further evidence. *Journal of Accounting Research* 24 (Autumn): 405–11.

———. 1988. Public accounting firms and the acquisition of nonaudit services by public and closely-held companies. *Auditing: A Journal of Practice & Theory* 8 (Fall): 63–77.

Public Oversight Board, Securities and Exchange Commission Practice Section, American Institute of Certified Public Accountants. 1986. *Public Perceptions of Management Advisory Services Performed by CPA Firms for Audit Clients.* New York: POB.

Scheiner, J. 1984. An empirical assessment of the impact of SEC nonaudit service disclosure requirements on independent auditors and their clients. *Journal of Accounting Research* 22 (Autumn): 789–97.

———, and J. Kiger. 1982. An empirical investigation of auditor involvement in nonaudit services. *Journal of Accounting Research* 20 (Autumn): 482–96.

Securities and Exchange Commission. 1978. *Accounting Series Release No. 250: Disclosure of Relationships with Independent Public Accountants.* Washington, D.C.: U.S. Government Printing Office (June).

———. 1981. *Accounting Series Release No. 296: Relationships Between Registrants and Independent Accountants.* Washington, D.C.: U.S. Government Printing Office (August).

———. 1982. *Accounting Series Release No. 304: Relationships Between Registrants and Independent Accountants.* Washington, D.C.: U.S. Government Printing Office (January).

Simunic, D. A. 1984. Auditing, consulting, and auditor independence. *Journal of Accounting Research* 22 (Autumn): 679–702.

Watts, R. L. 1977. Corporate financial statements: A product of the market and political processes. *Australian Journal of Management* 2 (April): 52–75.

———, and J. L. Zimmerman. 1983. Agency problems, auditing, and the theory of the firm: Some evidence. *Journal of Law and Economics* 26 (October): 613–33.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (May): 817–38.

SHORTER ARTICLES

1. Larry R. Davis, David N. Ricchiute, and Greg Trompeter
   Audit Effort, Audit Fees, and the Provision of Nonaudit Services to Audit Clients

2. Kenneth S. Lorek, Thomas F. Schaefer, and G. Lee Willinger
   Time-Series Properties and Predictive Ability of Funds Flow Variables

3. Frank A. Buckless and Robert L. Peace
   The Influence of the Source of Professional Standards on Juror Decision Making.

4. Jane F. Mutchler and Kurt Pany
   A Comment on "The Influence of the Source of Professional Standards on Juror Decision Making"

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 135–150

Audit Effort, Audit Fees, and the Provision of Nonaudit Services to Audit Clients

Larry R. Davis
University of Minnesota
David N. Ricchiute
University of Notre Dame
Greg Trompeter
Boston College

SYNOPSIS AND INTRODUCTION: In this article, we use audit-hour and billing-rate data supplied by a large public accounting firm to address the question, "Does providing audit clients with nonaudit services result in knowledge spillovers and audit production efficiencies that could produce economic rents for the auditor?" In prior analytical work, both Simunic (1984) and Beck et al. (1988) have argued that knowledge acquired while providing nonaudit services may "spill over" to the production of the audit, and thus generate production efficiencies. If audit production efficiencies lead to cost savings that are retained in whole or in part by the auditor (rather than passed on to the client), then economic rents accrue to the auditor, creating incentives for the auditor to resolve disputes in the client's favor.

Although several past studies suggest that the joint provision of audit and nonaudit services may give rise to knowledge spillovers that could lead to economic rents (Palmrose 1986; Simon 1985; Simunic 1984; Turpen 1990), they do not provide direct evidence that spillovers or rents exist, and the empirical results are mixed. For example, Abdel-khalik (1990, 320) reports that he was unable to detect interdependencies between audit and nonaudit fees, a direct contrast to the findings of Simunic and Palmrose. Further, Palmrose reports a positive relation between audit fees and the nonaudit fees paid to *non*incumbent firms, a finding that weakens the argu-

---

The authors express appreciation to the participating firm for providing data and to Larry Marsh, Tim Pearson, Rich Sheehan, Dan Simon, Ira Solomon, the referees, and workshop participants at the University of Notre Dame for helpful comments on earlier versions of this article.

Submitted March 1992.
Accepted September 1992.

ment for knowledge spillovers. Thus, as Solomon (1990, 328) points out, ". . . the impact of MAS (management advisory services) on audit pricing as well as who (i.e., the client or the auditor) benefits from knowledge spillovers (if they exist) remains an open and interesting question."

Our empirical analysis consists of three steps. First, we demonstrate the comparability of our sample to those in prior studies by fitting prior researchers' models to our data and replicating a finding Simunic (1984) interpreted as evidence of knowledge spillovers: a positive relation between audit and nonaudit fees. Second, we use a unique data set compiled from the participating firm's internal billing records, working papers, and audit planning memos to test for a positive relation between nonaudit services and audit effort. This test is motivated by Palmrose's (1986, 410) speculation that the higher audit fees paid by clients who also purchase nonaudit services may be driven by additional audit effort. We regress audit effort on nonaudit service fees (and Palmrose's control variables), partitioning nonaudit fees into three types: tax, accounting, and other. We use three measures of audit effort: unweighted audit hours, audit hours weighted by billing rate ratios, and audit hours weighted by billing rates. We find a weakly significant, positive relation between tax services and all three audit effort measures and between accounting-related consulting services and audit hours weighted by billing-rate ratios, which suggests that additonal effort is required for audits of clients who also purchase nonaudit services. With the assumption that the demand for auditing is inelastic (Beck et al. 1988, 52-54), these results do not support the existence of audit production efficiencies from knowledge spillovers.

Third, we test whether there is a positive relation between audit fees for a given level of audit effort and each of our three types of nonaudit service fees. This test is motivated by the possibility that, if the demand for auditing is elastic (Simunic 1984, 698), the observed increase in audit fees and effort could be driven by demand for more auditing by purchasers of nonaudit services (e.g., in substitution for internal control) as the result of auditors passing on cost savings from knowledge spillovers. If auditors are able to retain some of the cost savings, and thereby earn economic rents in the form of higher fees for a given level of audit effort, then our results should reveal a significant, positive relation between nonaudit and audit fees. However, we do not find a significant relation when we control for direct measures of audit effort that were not available to prior researchers.

These results suggest that, although purchasers of nonaudit services pay higher audit fees than nonpurchasers, the higher fees are associated with a proportional increase in audit effort, measured in this study as unweighted and weighted audit hours. These findings are inconsistent with one interpretation of prior research: that performing nonaudit services for audit clients may provide the auditor with incentives to compromise objectivity.

Key Words: *Auditing, Knowledge spillovers, Nonaudit services.*

Data Availability: *Requests for data used in this study may be made to the participating firm through the authors.*

THE remainder of this article is organized as follows: section I reviews prior research, section II describes the sample and data, section III reports the analysis and results, section IV discusses additional tests and limitations, and section V provides a summary and conclusion.

I. Prior Research

Over the past three decades, the effect of nonaudit services on auditor independence and objectivity has been addressed by Congressional committees (e.g., U.S. Senate 1977), special commissions (e.g., AICPA 1978), and more recently by the profession itself (e.g., AICPA 1991a; Arthur Andersen et al. 1991). A central concern to these groups and to academic researchers as well is whether auditors earn economic rents by providing nonaudit services to audit clients.

In audit fee research focusing on the joint provision of audit and nonaudit services, the general approach has been to regress audit fees on nonaudit fees, after controlling for differences across clients with financial statement measures and selected audit characteristics (Palmrose 1986; Simon 1985; Simunic 1984; Turpen 1990). These studies have reported a positive association between audit and nonaudit fees, and suggest that the audit fees paid by clients who also purchase nonaudit services are significantly higher than those paid by nonpurchasers.

Simunic (1984, 698), pointing to the interdependencies between audit and nonaudit services, interpreted the finding as evidence "... consistent with the hypothesis that the cost functions for MAS and auditing are significantly interdependent. Specifically ... the observed relationship would arise if the production of auditing generates knowledge useful in MAS production and/or the production of MAS reduces the marginal cost of auditing and audit demand is relatively elastic." This interpretation suggests that public accounting firms can provide audit (nonaudit) services more efficiently (i.e., with less effort) for nonaudit (audit) service purchasers and, thereby, earn economic rents.

However, Abdel-khalik (1990, 318) characterized the positive relation between audit and nonaudit fees as counterintuitive, noting that "... it is difficult to think of economic incentives that could exist *a priori* for clients to pay more for the joint acquisition of two products than for the sum of acquiring them separately." Several alternative explanations for the positive relation have been offered in the literature (see Abdel-khalik 1990; Palmrose 1986; Simunic 1984; Solomon 1990). For example, Palmrose found a positive relation between audit fees and nonaudit fees paid both to incumbent and to *non*incumbent audit firms (i.e., a firm other than the one performing the audit). Her finding for nonincumbent firms weakens the argument for spillovers, since it is unlikely that knowledge obtained by a nonincumbent audit firm from nonaudit services would spill over to the incumbent audit firm's production of the audit. She speculated that the positive relation for incumbent and nonincumbent firms could be the result of more effort being required to audit nonaudit service purchasers since, for example, nonaudit services may have audit implications (Palmrose 1986, 410). Thus, past research leaves unresolved the relationship among nonaudit services, audit effort, and audit fees. The objective of this study is to provide more direct evidence with respect to this issue by using proprietary data not reported previously in research on audit pricing (i.e., audit hours and staff billing rates).

II. Sample and Data

Sample data from a large public accounting firm were collected by members of the firm according to the authors' written instructions. Ten U.S. offices were asked to provide data on all publicly held audit clients and on an equal or greater number of privately held audit clients for the three most recent audits performed by the participating firm. From one to three years of usable data were returned for 98 clients. To avoid potential statistical problems related to the independence of observations, we included only one randomly chosen observation per client in the analyses.[1] Thus, the final sample consists of one observation for each of 98 clients.[2]

Data obtained from the firm's internal billing records consists of the actual fees charged for audit and nonaudit services, individual staff members' billing rates (see fn. 15), total audit hours by staff level, and out-of-pocket costs.[3] To test the robustness of our results using actual audit hours, we included two weighted measures of audit hours, both of which are based on the participating firm's hourly billing rates for individual staff members. The first weighted measure, total audit hours for an engagement normalized by the ratios between staff members' billing rates and the average billing rate in the sample (*RATIOHRS*), follows:

$$RATIOHRS = \frac{(\Sigma_i Hours_{ij} * Rate_i)}{\Sigma_i \Sigma_j (Rate_i * Hours_{ij}) / \Sigma_i \Sigma_j Hours_{ij}},$$

where,

Hours = the hours charged by a staff member to an engagement,
Rate = the firm's billing rate for one hour of a staff member's time,
i = an individual staff member, and
j = an individual audit engagement.

The next weighted measure, total audit hours for an engagement weighted by the staff members' billing rates plus out-of-pocket costs (*RATEHRS*), the participating firm's measure of audit costs for an engagement, is expressed as follows:

$$RATEHRS_j = (\Sigma_i Hours_{ij} * Rate_i) + Out\text{-}of\text{-}Pocket\ Costs_j,$$

where variables and subscripts are defined as above and out-of-pocket costs include travel, lodging, and meals.

Based on our discussions with firm personnel about the cycle of annual audit work, the firm provided audit-hour and fee data for services performed during the 12 months beginning eight months prior to the client's fiscal year-end. Fees for nonaudit services were collected for services performed during the same period.[4] In consideration of Palmrose's (1986) finding that the likelihood of knowledge spillovers may be a function of the type of nonaudit service provided, the firm partitioned nonaudit fees into tax (i.e., all tax consulting other than return preparation, such as consultation about the tax implications of financing alternatives), accounting (e.g., merger and acquisition advice, the design and installation of a new accounting system), and other services (i.e., nontax, nonaccounting services such as executive recruiting or business site selection) according to the firm's classification of services in its internal records.

---

[1] Tests for differences across audit years are reported in footnote 12.

[2] In total, we requested that the firm provide data for 225 clients; since data collection was labor intensive for the firm, we were reluctant to request a larger sample size. We received information on 176 clients. For inclusion in the final sample, we required that responses for each of the variables used in our models be provided. Thus, 44 clients were deleted because they were not billed separately for each type of service performed (i.e., the client received a single, non-itemized bill for all services). Twenty clients were deleted from the sample because of other missing billing or financial statement data. Finally, as we note in section IV, our test for a relationship between audit fees and the provision of nonaudit services may be sensitive to the time period over which we measure nonaudit services. This is of particular concern with new clients, and we therefore eliminated all observations for first-year engagements, further reducing our sample by 14 clients. The final sample of 98 clients consists of 38 public companies and 60 private companies.

[3] Throughout this article, the term "staff" refers to all auditors assigned to an engagement (e.g., engagement partner, manager, etc.), rather than to a particular rank (e.g., entry-level "staff" accountant).

To control for the potential effect of risk on audit fees, the firm collected the auditors' assessments of the risk associated with each engagement from audit planning memos (i.e., low, medium, or high risk). This risk measure is based primarily on the firm's assessment of the loss that will be incurred if an unqualified audit opinion is issued inappropriately. Although risk measures assessed during the planning stage of an engagement are arguably subjective, we believe they more closely reflect auditors' actual perceptions of risk, and therefore the effect of risk on audit fees, than the surrogates used in prior research. Descriptive statistics for these variables and for assets (a size measure) are in table 1.[5]

To permit us to fit prior researchers' models to our data, the firm collected financial statement measures and selected characteristics of the audit from copies of clients' annual reports retained by the firm and from audit working papers. These variables included net receivables, inventory, total assets, opinion type, number of locations visited by the auditor, number of reports issued by the auditor, primary SIC code,[6] ownership (public or private), number of subsidiaries, and number of industries in which business is conducted.

III. Analysis and Results

Replication of Past Findings

Table 2 presents the results of regression analyses to fit the models used in previous research to our data. Although several variables are nonsignificant[7] and our adjusted

---

[4] In section IV, we test the sensitivity of our results to the period over which nonaudit services are measured.

[5] Observe in table 1 that the mean total asset size of our sample of 64 nonaudit purchasers is larger than that of our sample of 34 nonpurchasers ($56.8 vs. $12.3 million). To address the structural stability of our models (using, in turn, each of our three measures of audit effort), we performed Chow tests for our sample of nonaudit purchasers and nonpurchasers (Kennedy 1985, 87–89). The overall F-test indicates that the regression coefficients are stable across subsamples. We also performed a dummy variable version of the Chow test that allows tests for a specific coefficient or a subset of coefficients (Kennedy 1985, 186). In this test, we focused on the input measure, since it is most directly correlated with size. This test requires a new variable, *DUMMY*INPUT*, in the regression equation that takes on a value equal to the input measure if the observation represents a client who purchased nonaudit services, and a value of zero for nonpurchasers. Results of the tests indicate that the regression coefficients in tables 2, 3, and 4 are robust across our samples.

[6] Using two-digit SIC codes, we classified companies in the following industries: construction, financial institutions, manufacturing, mining, natural resources, retail, service, transportation, and wholesale.

[7] For example, in comparison with Simunic's (1984) and Palmrose's (1986) small-companies subsamples (their subsamples that are most comparable to ours), the following variables are significant in their analyses but not in ours. Simunic: audit opinion type, number of subsidiaries; Palmrose: number of audit locations, publicly/privately held. None of the significant variables in our table 2 replications of Simunic and Palmrose were nonsignificant for the smaller-companies analyses reported in their papers.

Table 1
Descriptive Statistics

	Purchasers of Nonaudit Services	*Nonpurchasers*
Total assets (000)	\$56,836.6 (132,448.3)	\$12,349.7 (15,311.7)
Audit hours	727 (591)	438 (349)
Hours weighted by billing rate ratios	730 (575)	431 (341)
Hours weighted by billing rates (000)	\$63.8 (50.8)	\$38.5 (30.8)
Audit fees (000)	\$55.0 (46.5)	\$28.9 (24.4)
Noncompliance tax fees (000)	\$8.5 (14.7)	—
Accounting-related nonaudit fees (000)	\$14.7 (19.7)	—
Other nonaudit fees (000)	\$1.2 (6.7)	—
High risk (percentage)	14.1	2.9
Medium risk (percentage)	51.6	41.1
Low risk (percentage)	34.3	56.0
Number of observations	64	34

R^2s (0.22, 0.24, 0.53, and 0.52) are lower than those reported in the previous four studies listed in table 2 (0.36, 0.54, 0.89, and 0.81, respectively), our results are generally consistent. Assets, a proxy for client size, are highly correlated with fees, as are other variables used in the models to proxy for financial or audit complexity and risk (e.g., number of subsidiaries and opinion type). Of greatest interest is the finding that nonaudit fees are positively associated with audit fees. The only exception is the nonsignificant coefficient for other nonaudit fees (i.e., nontax, nonaccounting services) in Palmrose's (1986) model; however, this is consistent with Palmrose's findings for smaller companies, the subsample in her study that is most comparable to ours. These findings indicate that our sample is comparable to samples used in prior research, since the positive relation between audit and nonaudit fees, interpreted in prior research as evidence of knowledge spillovers, is also present in our sample. In the next section we describe our analyses designed to test for a positive relation between nonaudit services and audit effort.

Tests for a Relation Between Nonaudit Services and Audit Effort

The positive relation between audit and nonaudit fees reported in table 2 and in previous research (Palmrose 1986; Simon 1985; Simunic 1984; Turpen 1990) indicates that clients who spend more for nonaudit services may pay higher audit fees. In view of

Table 2
Replication of Positive Relation Between Audit and Nonaudit Fees by Regressing Audit Fees on Prior Researchers' Variables
(n=98)

	Simunic[a] (1984)	*Simon*[b] (1985)	*Palmrose*[b] (1986)	*Turpen*[b] (1990)
Control variables:				
Intercept	9.09 (3.36)***	4.82 (1.68)**	4.97 (6.04)***	5.57 (7.62)***
Assets	—	—	0.29 (6.07)***	0.26 (5.61)***
Number of audit locations	—	—	0.09 (0.83)	—
Audit opinion type	0.78 (0.88)	0.70 (0.81)	0.14 (2.29)**	0.11 (1.97)**
Industry indicators	—	—	[c]	[c]
Publicly/privately held	—	—	0.03 (0.61)	—
Number of reports issued by auditor	—	—	0.47 (2.96)***	—
Accounts receivable	−0.86 (−0.19)	−1.29 (−0.30)	—	—
Inventory	−1.92 (−0.43)	−1.49 (−0.34)	—	—
Number of subsidiaries	−0.33 (−0.45)	−0.14 (−1.37)*	—	0.15 (2.97)***
Number of segments	4.00 (2.99)***	4.38 (3.33)***	—	—
Financial institution	0.10 (0.03)	—	—	—
Current assets	—	—	—	0.18 (0.81)
Experimental variables:				
Total nonaudit fees	0.65 (4.39)**	0.66 (4.54)***	—	0.04 (2.89)***
Noncompliance tax-related nonaudit fees	—	—	0.03 (1.65)*	—
Accounting-related nonaudit fees	—	—	0.02 (1.49)*	—
Other nonaudit fees	—	—	0.02 (0.95)	—
Adjusted R^2	0.22	0.24	0.53	0.52

Note: The *t*-values for each parameter estimate appear in parentheses below the estimate.

[a] Audit fees and nonaudit fees are deflated by the square root of total assets. Inventory and accounts receivable are deflated by total assets.

[b] Audit fees, assets, and nonaudit fees are set equal to their natural logs.

[c] Coefficients of industry indicator variables based on two-digit SIC codes are not significant.

* Significant at $p \leq 0.10$ (one-tailed test).

** Significant at $p < 0.05$ (one-tailed test).

*** Significant at $p < 0.01$ (one-tailed test).

Palmrose's (1986, 410) speculation, we test whether the positive relation is driven by auditors expending more effort to audit the financial statements of clients who spend more for nonaudit services. To do this, we regress audit effort on the three nonaudit service variables—tax, accounting, and other—and the control variables Palmrose (1986) used to predict audit fees. We use Palmrose's (1986) control variables because she justifies her model primarily on the basis of the relationship between audit effort and her independent variables, and she also partitioned nonaudit services into tax, accounting, and other. Each of the three alternative measures of audit effort introduced in section II (audit hours, audit hours weighted by billing rate ratios, and audit hours weighted by billing rates) is used in a separate regression to test the robustness of our findings. In our regressions, a positive relation between the purchase of nonaudit services and audit effort would be evidence of auditors' expending more audit effort.

As was true for the prediction of audit fees (table 2), the results reported in table 3 indicate that measures of client size (assets) and audit complexity (e.g., number of locations visited) are positively associated with audit effort. Most important, the coefficients on all nonaudit services are positive, the coefficient on tax services is positive and weakly significant for all three audit effort measures, and the accounting-related services measure is positive and weakly significant for the *RATIOHRS* measure of audit effort.[8] These results suggest that those nonaudit services most closely related to a client's financial activities (i.e., tax and accounting-related nonaudit services) are positively related to audit effort. That is, contrary to the intuitive expectation that knowledge spillovers would allow the auditor to perform the audit with less effort, we find evidence that clients who spend more for nonaudit services typically require more audit effort.

One possible explanation for this finding is that nonaudit services may have audit implications (Palmrose 1986, 410). For example, consider a client who purchases tax advice that leads to the use of leases rather than loans to finance asset acquisitions, or who purchases accounting services that lead to a new accounting information system. In both cases, the provision of nonaudit services would likely require the auditor to gather a larger quantity of audit evidence because leases may increase financial complexity, and a new accounting system creates the need to document the new internal control structure (AICPA 1991b, AU sec. 319).

The results in table 3 contradict the existence of audit production efficiencies from knowledge spillovers, with the assumption that the demand for auditing is *in*elastic (Beck et al. 1988, 52–54). However, if the demand for auditing is elastic (Simunic 1984, 698), then the results in table 3 are not necessarily inconsistent with knowlege spillovers. That is, if auditors pass on cost savings from spillovers, then the increase in audit effort could be driven by nonaudit purchasers' acquiring more auditing (Simunic 1984). For example, some clients may be compelled by audit cost savings to expend less for internal control, electing instead to substitute the purchase of auditing services from accounting firms. As a consequence, factors such as a weaker internal control structure would necessitate more audit hours. In this scenario, auditors would earn economic rents from nonaudit service purchasers in the form of higher fees for a given level of effort if audit cost savings from spillovers are retained in part by the auditor. We address this next.

---

[8] In addition to the regression results reported in table 3, we also pooled the three nonaudit service variables (*TAX*, *ACC*, *OTHER*) and found that total combined nonaudit service fees is significant for all three measures of effort at p levels of less than 0.02 (one-tailed test).

Table 3

Tests for Relation Between Nonaudit Services and Quantity of Effort Used in the Production of an Audit

(n=98)

	ℓn(audit hours)	*ℓn(audit hours weighted by billing rate ratios)*	*ℓn(audit hours weighted by billing rates)*
Control variables:			
Intercept	1.63 (2.24)**	1.54 (2.01)**	5.71 (7.62)***
Assets	0.24 (5.77)***	0.25 (5.86)***	0.27 (6.21)***
Number of audit locations	0.17 (1.80)**	0.16 (1.63)*	0.13 (1.41)*
Audit opinion type	0.07 (1.39)*	0.09 (1.73)**	0.11 (2.08)**
Industry indicators	[a]	[a]	[a]
Publicly/privately held	−0.02 (−0.41)	−0.004 (−0.09)	−0.001 (−0.03)
Number of reports issued by auditor	0.46 (3.26)***	0.41 (2.88)***	0.41 (2.85)***
Experimental variables:			
Noncompliance tax-related nonaudit fees	0.02 (1.47)*	0.02 (1.35)*	0.03 (1.53)*
Accounting-related nonaudit fees	0.02 (1.26)	0.02 (1.62)*	0.02 (1.22)
Other nonaudit fees	0.01 (0.78)	0.01 (0.72)	0.01 (0.54)
Adjusted R^2	0.54	0.54	0.54

Note: The *t*-values for each parameter estimate appear in parentheses below the estimate.

[a] Coefficients of industry indicator variables based on two-digit SIC codes are not significant.

* Significant at $p<0.10$ (one-tailed test).
** Significant at $p<0.05$ (one-tailed test).
*** Significant at $p<0.01$ (one-tailed test).

Tests for a Relation Between Nonaudit Services and Audit Fees

In this section, we test for a positive relation between the provision of nonaudit services and audit fees for a given level of audit effort.[9] We regress total audit fees on audit effort, risk, and the primary experimental variables of interest: nonaudit fees for tax, accounting, and other consulting services.

---

[9] Although a positive relation between nonaudit services and audit fees would suggest that the accounting firm realized economic rents, no relation would not rule out economic rents. For example, the auditor could earn rents through increased tenure or through nonaudit service fees. Thus, the tests reported are a sufficient but not a necessary condition for economic rents.

To control for heteroscedasticity,[10] we use weighted least squares (WLS) regression (Judge et al. 1985).[11] WLS produces unbiased estimates of the regression coefficients that can be interpreted in the same manner as OLS estimates. However, because WLS weightings could lead to inflated R^2s, we also report ordinary least squares (OLS) results for the models below as a benchmark for comparison (see table 4). The use of WLS involves the following three-step procedure for each model: (1) with OLS, audit fees are regressed on the effort measure; (2) the absolute values of the residuals from step 1 are regressed on the effort measure; and (3) for each observation, the variables are weighted by the reciprocal of the predicted residual from step 2. Formal statements of the models are as follows:

$$\begin{aligned} FEES_w = {} & B_0 + B_1 HRS_w + B_2 HRISK_w + B_3 MRISK_w + B_4 TAX_w \\ & + B_5 ACC_w + B_5 OTHER_w \end{aligned} \tag{1}$$

$$\begin{aligned} FEES_w = {} & B_0 + B_1 RATIOHRS_w + B_2 HRISK_w + B_3 MRISK_w \\ & + B_4 TAX_w + B_5 ACC_w + B_6 OTHER_W \end{aligned} \tag{2}$$

$$\begin{aligned} FEES_w = {} & B_0 + B_1 RATEHRS_w + B_2 HRISK_w + B_3 MRISK_w \\ & + \mathrm{B}_4 TAX_w + B_5 ACC_w + B_6 OTHER_w \end{aligned} \tag{3}$$

with the "*w*" subscript denoting a weighted variable as described above. Variables are defined as follows:

FEES = the total audit fee paid to the auditor,
HRS = the direct labor hours used in the production of the audit,
RATIOHRS = *HRS* weighted by the ratio of the billing rates of the individuals providing the labor to the average billing rate for all audit hours in our sample,
RATEHRS = *HRS* weighted by the billing rates for the individuals providing the labor, plus any out-of-pocket costs associated with the audit,
HRISK = an indicator variable coded 1 if the auditor's assessment of risk was high, 0 otherwise,
MRISK = an indicator variable coded 1 if the auditor's assessment of risk was moderate, 0 otherwise,
TAX = total fees paid to the audit firm for tax consulting services,
ACC = total fees paid to the audit firm for accounting related consulting services, and
OTHER = total fees paid to the audit firm for non-tax, non-accounting related consulting services.

---

[10] Scatter plots of the data (not reported) indicate that the relation between each of the three effort measures and audit fees is linear but that the absolute values of the regression residuals are positively correlated with the size of the dependent variable's predicted values. A Goldfeld-Quandt test for heteroscedasticity was significant at $p < 0.001$ for each of our three effort measures.

[11] Prior audit-fee research predicts audit fees as a function of asset size and accordingly uses either log or square root transformations to correct for heteroscedasticity. As depicted in data plots (not reported), the audit fees in this study increase with asset size at a decreasing rate, thus making such transformations appropriate when modeling audit fees as a function of asset size. However, our models predict audit fees as a function of audit effort, and, as one would expect, scatter plots (not reported) indicate a strong linear relation between fees and effort (i.e., as a function of audit effort, the actual fee increases at an approximately constant rate). For this reason, weighted least squares (WLS) is more appropriate here for controlling heteroscedasticity than are transformation techniques such as log or square root transformations.

Table 4
Weighted Least Squares and Ordinary Least Squares Regressions of Audit Fees on Audit Effort, Nonaudit Fees, and Risk
(n=98)

	Model 1		Model 2		Model 3	
	WLS[a]	*OLS*[b]	*WLS*[c]	*OLS*[d]	*WLS*[e]	*OLS*[f]
Control variables:						
INTERCEPT	371.14 (0.19)	4.36 (13.71)***	175.81 (0.10)	4.32 (14.24)***	−1,834.44 (−2.05)**	−0.58 (−1.93)*
HRS	68.82 (14.67)***	0.96 (17.15)***	—	—	—	—
RATIOHRS	—	—	71.45 (15.24)***	0.97 (18.16)***	—	—
RATEHRS	—	—	—	—	0.86 (34.01)***	1.03 (34.56)***
HRISK	5,848.80 (1.00)	0.18 (1.27)*	7,399.64 (1.30)*	0.20 (1.52)*	4,311.90 (1.57)*	0.08 (1.05)
MRISK	34.94 (0.01)	0.04 (0.46)	−1,040.00 (−0.42)	0.01 (0.10)	172.04 (0.14)	0.02 (0.44)

Continued

Table 4—*Continued*

Weighted Least Squares and Ordinary Least Squares Regressions of Audit Fees on Audit Effort, Nonaudit Fees and Risk

(n=98)

	Model 1		*Model 2*		*Model 3*	
	WLS[a]	*OLS*[b]	*WLS*[c]	*OLS*[d]	*WLS*[e]	*OLS*[f]
Experimental variables:						
TAX	0.15 (0.93)	0.01 (0.71)	0.11 (0.69)	0.01 (1.21)	−0.02 (−0.20)	<0.01 (0.62)
ACC	0.09 (0.71)	0.01 (0.89)	0.01 (0.10)	−0.01 (−0.02)	−0.03 (−0.52)	<0.01 (0.94)
OTHER	0.11 (0.46)	<0.01 (0.10)	<0.01 (0.01)	<0.01 (0.04)	0.10 (0.90)	0.01 (0.83)
Adjusted R^2	0.79	0.83	0.81	0.85	0.95	0.95

Note: The *t*-values for each parameter estimate appear in parentheses below the estimate.

[a] $FEES_w = B_0 + B_1 HRS_w + B_2 HRISK_w + B_3 MRISK_w + B_4 TAX_w + B_5 ACC_w + B_6 OTHER_w$

[b] $LFEE = B_0 + B_1 \ell n(HRS) + B_2 HRISK + B_3 MRISK + B_4 \ell n(TAX) + B_5 \ell n(ACC) + B_6 \ell n(OTHER)$

[c] $FEES_w = B_0 + B_1 RATIOHRS_w + B_2 HRISK_w + B_3 MRISK_w + B_4 TAX_w + B_5 ACC_w + B_6 OTHER_w$

[d] $LFEE = B_0 + B_1 \ell n(RATIOHRS) + B_2 HRISK + B_3 MRISK + B_4 \ell n(TAX) + B_5 \ell n(ACC) + B_6 \ell n(OTHER)$

[e] $FEES_w = B_0 + B_1 RATEHRS_w + B_2 HRISK_w + B_3 MRISK_w + B_4 TAX_w + B_5 ACC_w + B_6 OTHER_w$

[f] $LFEE = B_0 + B_1 \ell n(RATEHRS) + B_2 HRISK + B_3 MRISK + B_4 \ell n(TAX) + B_3 \ell n(ACC) + B_6 \ell n(OTHER)$

* Significant at $p<0.10$ (one-tailed test).
** Significant at $p<0.05$ (one-tailed test).
*** Significant at $p<0.01$ (one-tailed test).

The regression results are presented in table 4. The WLS (OLS) R^2s for all three models—0.79 (0.83), 0.81 (0.85), and 0.95 (0.95), respectively—indicate a good linear fit and that the regression models explain a large proportion of the cross-sectional variation in audit fees.[12] As expected, the coefficients for audit effort are positive and statistically significant in each model. The WLS results for models 1 and 2 indicate that, on average, $69 of revenue is realized per unweighted audit hour and that $71 of revenue is realized per weighted audit hour. The results for model 3 indicate that, on average, the firm realized as revenue 86 percent of audit hours weighted by billing rates.

The most important finding is the nonsignificant coefficients on the nonaudit services variables (*TAX, ACC, OTHER*) in all three models.[13] This indicates that, controlling for audit effort and risk, these nonaudit service variables are not related to audit fees.[14] That is, relative to the effort auditors expend on an audit, the level of nonaudit fees received from audit clients is not associated with a significantly higher (or lower) audit fee. Thus, our results do not support the existence of economic rents accruing to auditors as the result of knowledge obtained from performing nonaudit services. This suggests either that knowledge spillovers from nonaudit services do not lead to audit production efficiencies or, alternatively, that the benefits of efficiencies are passed on to clients.

To test the sensitivity of our results, we fit expanded models to our sample, adding indicator vaiables for office location (city), industry, and ownership structure (public vs. private companies). Adding office location to the analysis is motivated by Abdel-khalik's (1990, 297) argument that a firm's monopoly power within a city may reduce its incentives to pass on to clients the cost savings attributable to knowledge spillovers. Adding industry and ownership structure is motivated by research that indicates that audit fees may vary across industries and as a function of ownership structure (see, e.g., Palmrose 1986; Simunic 1984). Results for the expanded models 1, 2, and 3 reveal that all three nonaudit services (*TAX, ACC, OTHER*) are nonsignificant, indicating that the addition of office location, industry, and ownership structure does not significantly affect the results reported in table 4.

We also performed tests using *FEES* divided by *RATEHRS* as the dependent variable to represent an engagement profitability measure referred to by the firm as the "realization rate."[15] If the joint provision of audit and nonaudit services results in knowledge spillovers and production efficiencies beneficial to the auditor, then we would expect a positive relation between nonaudit service fees and realization rates. To test for a relation between the purchase of nonaudit services and audit fees for a given level of effort, we regressed realization rates on nonaudit service fees and risk. In conformance with the results reported in table 4, we fail to find a significant relation, additional evidence that the joint provision of services does not lead to economic rents in the form of higher fees for a given level of audit effort.

---

[12] Tests indicated that our primary results are not affected by the fiscal year that was audited or the time of year (i.e., month) when work was performed. Tests for fiscal-year effects were performed by incorporating categorical variables into the regression models. Tests for time-of-year effects were performed by including in the models a variable defined as the total audit hours charged by an office for all engagements during the month when the audit report was signed, divided by the total audit hours charged to all audit engagements by the office for the fiscal year as a surrogate for how "busy" a month is relative to all other months in the year. The results of these tests indicate that the audit fees received for a given level of effort do not vary by time of year or fiscal year, and do not affect our primary tests of the relation between audit and nonaudit fees. For the sake of expositional clarity, we report only the reduced models.

[13] Diagnostic procedures (Belsley et. al. 1980) indicate that the lack of significant coefficients on nonaudit service fees reported in table 4 is not the result of collinearity among the variables in our models.

[14] In addition to the tests reported in table 4, we also pooled the three nonaudit service variables (*TAX, ACC,* and *OTHER*) and found that total combined nonaudit services is not significant in any of the three models.

[15] For the participating firm, a staff member's billing rate represents the maximum amount the firm would expect to charge for one hour of that employee's labor. However, actual audit fees are generally less than *RATEHRS* (actual hours × billing rates) because of such factors as competition from other firms and slack time or periods of relatively low audit activity (e.g., the summer months in a practice office dominated by December 31 year-end clients). Thus, realization rates commonly have a value less than 1. For example, in our sample, WLS realization rates ranged from 0.51 to 1 (mean=0.86, quartile values of 0.7, 0.86, 0. 92 and 1).

IV. Additional Tests and Limitations

In our analyses, nonaudit fees are reported for services rendered during the same time period in which the audit is normally conducted. While Simunic (1984) suggests that spillovers may result from providing both audit and nonaudit services in the same year, Beck et al. (1988) recognize that spillovers may result from services performed in previous years. We performed three additional analyses to test the sensitivity of our findings to the period over which nonaudit services are measured.

First, for the 52 clients for which we have two consecutive years of data, we computed the correlations between the nonaudit services purchased by clients from one year to the next. According to Palmrose (1986) and our replication of her findings (see table 2) for smaller companies like those in our sample, tax and accounting-related nonaudit services are the two types of nonaudit services most likely to be associated with relatively higher audit fees. The year-to-year correlations for these two types of services were 0.78 and 0.68, respectively. This suggests that clients who purchase these services tend to do so consistently over time and that fees for these nonaudit services in one 12-month period are a good surrogate for fees in subsequent 12-month periods.[16]

Second, for the same 52 clients, we modified the analysis reported in table 4, by including in the regression models the nonaudit fees for the 12 months preceding the period when the audit is performed. These fees were included in the models as new variables and, alternatively, as increases in the nonaudit fees used in our primary analysis. In conformance with the results reported, we do not find a significant relation between nonaudit and audit fees charged for a given level of effort.

Third, we modified the analysis in table 4 by using only the most recent of the years for which data were collected (rather than randomly choosing one year) and classifying as nonpurchasers only those clients who did not purchase nonaudit services during any of the years for which data were collected. We include in the regression models the nonaudit fees for the two 12-month periods preceding the audit period. Again, in conformance with the results reported, we find no significant relation between nonaudit fees and the audit fees charged for a given level of effort. In combination, these three tests suggest that our results are not affected significantly by the manner in which we measured nonaudit fees.

As Solomon (1990, 325) points out, audit partners' compensation schemes and their influence relative to consulting partners may provide incentives for them to categorize fees arbitrarily as audit or nonaudit. For example, in a period when total fees have fallen from a prior period, an audit partner who acquired a client originally may arbitrarily categorize more of the fee as deriving from audit services. Although we have neither empirical nor anecdotal evidence of such practice, our results are subject to the limitation that nonaudit fees may be misclassified.

---

[16] For nontax, nonaccounting services, the year-to-year correlation for fees is 0.23. In conformance with Palmrose's (1986) findings for smaller companies, the results reported in table 2 suggest that, for the clients in our sample, this type of service is unrelated to factors that lead to relatively higher audit fees.

As indicated previously, the mean total asset size of our sample of 64 nonaudit service purchasers ($56.8 million) is larger than that of our 34 nonpurchasers ($12.3 million). Although various tests (see fn. 5) indicate that the regression coefficients in tables 2, 3, and 4 are robust across samples, our results are subject to the limitation that size may be partially driving the weakly significant coefficients for tax and accounting-related nonaudit services in table 3. Our results are also subject to the limitations that we obtained data from only one public accounting firm and that our sample consists of small companies, most of which are private (see fn. 2). Differences between the audit technologies of competing firms and firms' pricing structures for large versus small companies and for public versus private companies may affect the extent to which the joint provision of services produces knowledge spillovers and production efficiencies for the auditor. Thus, the generalizability of our findings to other firms and larger companies is an open question.

The failure to find a relation between audit fees for a given level of audit effort and nonaudit fees raises the question of the adequacy of our statistical power. However, the high explanatory power of our models, the small magnitude of the *t*-statistics for nonaudit fees in our models, and our ability to replicate past findings indicate that any production efficiencies that do arise from knowledge spillovers are—for our sample—likely quite small.

V. Summary and Conclusion

Simunic (1984) reported a positive association between audit and nonaudit fees, and suggested that this result could be explained either by knowledge spillovers or by systematic differences between nonaudit purchasers and nonpurchasers. Palmrose's (1986) findings of a positive relation between audit fees and the nonaudit fees paid to nonincumbent firms weaken the argument for knowledge spillovers but, lacking data on audit effort, she was unable to provide direct evidence that spillovers, production efficiencies, or economic rents do not exist. She speculated that nonaudit service purchasers may require more audit effort, and the results reported in this article are consistent with that reasoning.

We replicate past findings of a positive relation between audit and nonaudit fees by fitting prior researchers' models to our data (table 2), and we find that nonaudit purchasers require more audit effort (table 3). With the assumption that the demand for auditing is inelastic, the results in table 3 are inconsistent with audit production efficiencies arising from knowledge spillovers, a condition that could lead to economic rents for the auditor. Conversely, with the assumption that the demand for auditing is elastic, the results in table 3 do not rule out spillovers. However, the results in table 4 provide no evidence that spillovers, if they exist, allow auditors to receive higher fees for a given level of audit effort from purchasers of nonaudit services than they receive from nonpurchasers. The results reported in this article provide no empirical evidence for the argument that providing nonaudit services for audit clients creates circumstances that may lead auditors to compromise their objectivity. Further research is needed to investigate other ways in which auditors may benefit from the joint provision of services, such as the effects of auditor tenure on client-firm bonding and economic rents earned through *non*audit services.

References

Abdel-khalik, A. R. 1990. The jointness of audit fees and demand for MAS: A self-selection analysis. *Contemporary Accounting Research* 7 (Spring): 295–322.

American Institute of Certified Public Accountants. 1978. *Commission on Auditors' Responsibilities: Report, Conclusions, and Recommendations.* New York: AICPA.

———. 1991a. *Statement on Consulting Services, Definitions and Standards.* New York: AICPA.

———. 1991b. *Codification of Statements on Auditing Standards.* New York: AICPA.

Arthur Andersen, Coopers & Lybrand, Deloitte and Touche, Ernst & Young, KPMG, Peat Marwick, and Price Waterhouse. 1991. *The Public Accounting Profession: Meeting the Needs of a Changing World.*

Beck, P. J., T. J. Frecka, and I. Solomon. 1988. A model of the market for MAS and audit services: Knowledge spillovers and auditor-auditee bonding. *Journal of Accounting Literature* 7: 50–64.

Belsley, W. R., F. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York: John Wiley & Sons.

Judge, G., W. Griffiths, R. C. Hill, H. Lutkepohl, and T. Lee. 1985. *The Theory and Practice of Econometrics.* New York: John Wiley & Sons.

Kennedy, P. 1985. *A Guide to Econometrics.* Cambridge, MA: MIT Press.

Palmrose, Z. 1986. The effect of nonaudit services on the pricing of audit services. *Journal of Accounting Research* 24 (Autumn): 405–11.

Simon, D. 1985. The audit services market: Additional empirical evidence. *Auditing: A Journal of Practice & Theory* 5 (Fall): 71–78.

Simunic, D. 1984. Auditing, consulting, and auditor independence. *Journal of Accounting Research* 22 (Autumn): 679–702.

Solomon, I. 1990. Discussion of the jointness of audit fees and demand for MAS: A self-selection analysis. *Contemporary Accounting Research* 7 (Spring): 323–28.

Turpen, R. A. 1990. Differential pricing on auditors' initial engagements: Further evidence. *Auditing: A Journal of Practice & Theory* 9 (Spring): 60–76.

U.S. Senate. 1977. Subcommittee on Reports, Accounting, and Management of the Commission on Government Operations. *The Accounting Establishment: A Staff Study.* Washington, D.C.: U.S. Government Printing Office.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 151–163

Time-Series Properties and Predictive Ability of Funds Flow Variables

Kenneth S. Lorek
Thomas F. Schaefer
Florida State University
G. Lee Willinger
University of Oklahoma

SYNOPSIS AND INTRODUCTION: An accurate description of the process that generates measures of funds flow (cash flow and working capital from operations) has potential importance in a variety of decision contexts. In particular, the issue of cash flow prediction has been central to standard setters. According to Statement of Financial Accounting Concepts No. 1 (FASB 1978, par. 37) "financial reporting should provide information to help investors, creditors, and others assess the amount, timing, and uncertainty of prospective net cash inflows to the related enterprise."

Our first objective in this study is to provide descriptive evidence that documents the statistical patterns (e.g., seasonality, autocorrelation) of the cash flow and working capital series for a sample of firms. Although the accounting literature is replete with evidence on the time-series properties of earnings numbers, much less systematic evidence on such properties for cash flow and working capital series is presently available. Our results suggest that the time-series behavior of the cash flow series is in marked contrast to the models typically employed for accounting net income. We also provide evidence that the working capital series behaves similarly to net income.

Funds flow is important in the investigation of security price effects, and the association of various measures of funds flow with security returns has been studied by Bernard and Stober (1989), Bowen et al. (1987), Schaefer and Kennelley (1986), and Wilson (1987), among others. These studies have employed cross-sectional expectation models for funds flow

We appreciate the comments of the referees and G. Peter Wilson as well as those of participants in the Florida State University and University of Notre Dame accounting research colloquia and the Southeast AAA doctoral consortium.

Submitted June 1990.
Accepted August 1992.

measures that restrict coefficients to be the same across firms. We assess the effect of such restrictions on the predictive ability of these models by comparing them with univariate time-series models that permit firm-specific estimation of coefficients. Accordingly, our second objective is to assess the accuracy of forecasts of funds flow variables generated by univariate time-series models versus those obtained from the multivariate cross-sectional models used in prior research (see, e.g., Bernard and Stober 1989; Wilson 1986, 1987).

This study provides new evidence on the time-series properties of cash flow and working capital series. The empirical results indicate that the statistical patterns of the cash flow series stand in marked contrast to the well-documented characteristics of quarterly earnings data. Cash flow series are modeled parsimoniously by purely *seasonal* time-series models. Specifically, we provide descriptive and predictive evidence supportive of the (000) × (100) seasonal Autoregressive Integrated Moving Average (ARIMA) model as a candidate model for predicting cash flow. This model outperforms the multivariate cross-sectional models used in prior research in out-of-sample predictive ability tests. We also present evidence that working capital from operations exhibits time-series behavior virtually identical to that of accounting earnings. This results in identification of ARIMA expectation models quite similar to those popularized for quarterly earnings. Such univariate ARIMA models for working capital dominated cross-sectional regression models in tests of predictive ability.

Key Words: *Cash flow, Working capital, Time-series models.*

Data Availability: *The data used in this study are available on the COMPUSTAT and CRSP data bases.*

THE article is organized as follows: An initial background section is followed by a description of the funds flow measures and sample selection criteria. Next, sections on descriptive evidence, the findings on predictive ability, and security return effects are provided. Finally, concluding remarks are presented.

I. Background

Despite the wide variety of potential uses of cash flow data in prediction contexts (Bowen et al. 1986), little is known about the time-series properties of quarterly cash flow and other funds flow measures. Bowen et al. (1986) report that a relatively high correlation exists between percentage changes in a traditional cash flow measure such as working capital per share and percentage changes in earnings. Nevertheless, relatively low correlation exists between percentage changes in a more refined cash flow measure, cash flow from operations, and percentage changes in earnings. This suggests that the employment of separately identified structures in modeling different cash flow series may be appropriate.

We use statistical goodness-of-fit criteria to identify univariate time-series models in addition to the cross-sectional expectation models used by Wilson (1987) and Bernard and Stober (1989).[1] Such cross-sectional expectation models for cash flow and working capital from operations may have several drawbacks when employed in this context. First, such models assume implicitly that the time-series properties of quarterly funds flow measures and net earnings are similar in structure across measures and identical in parameter values across firms. Our empirical results underscore the marked difference in the time-series properties of cash flow from operations versus those displayed by net earnings. Therefore, the employment of distinct prediction models for such series may be appropriate.

Second, lagged values of several variables, including sales, cash from operations, current and noncurrent accruals, are employed in cross-sectional models without explicit justification for the lag structure employed ($t-1$, $t-2$, $t-3$, $t-4$, etc.). Third, inspection of the set of independent variables used in the cross-sectional regressions reveals the employment of contemporaneous values of sales and net earnings relative to the dependent variables for all sample firms. Wilson (1987, 307) provides empirical evidence that inclusion of such variables in a cash flow multivariate model has a negligible effect on overall explanatory power.[2] These results suggest that a simple univariate model employing historical data may perform relatively well and may represent a viable alternative to the multivariate models.

Funds Flow Measures and Sample Selection Criteria

Two separate measures of funds flow, working capital flow from operations (WCF) and cash flow from operations (CF) are calculated on a per-share basis by using information from the annual and quarterly COMPUSTAT data bases. We constructed a time-series of quarterly observations for both measures beginning in the first quarter of 1976 and ending in the fourth quarter of 1985 for 109 sample firms. We used the first 36 observations of data to identify and estimate the parameters of the time-series expectation models. We withheld the four quarterly observations in 1985 for predictive ability tests. Henceforth, we refer to this group of firms as the 1985 sample. We then obtained four additional quarters of information in 1986, which resulted in a reduced sample of 80 firms. For these firms, we used the first 40 quarters of data to estimate the parameters of the expectation models and withheld the four quarters of data in 1986 for assessment of predictive ability. These 80 firms are referred to subsequently as the 1986 sample. Table 1 details the specific sample selection criteria for both samples. Each year's final sample size was influenced by the availability of component income statement and balance sheet accounts on quarterly COMPUSTAT, interest expense data on annual COMPUSTAT, and the employment of a December fiscal year-end filter.[3]

Because directly reported amounts for measures of funds flow were not available across the time period we analyze, we constructed values for the variables by aggregating their components from quarterly reports available through COMPUSTAT. For comparability, our aggregation technique paralleled the procedures used by Bernard and

---

[1] We also assess two time-series multiple regression models that employ alternative earnings figures as independent variables on the cash flow series. We are indebted to an anonymous reviewer for suggesting these models.

[2] Wilson (1987) reports that inclusion of currently revealed earnings and revenue numbers accounts for only 1.2 percent of the model's adjusted R^2 of 47.3 percent. Apparently, the lagged variables provided almost all of the information needed for estimation.

[3] The December fiscal year-end requirement assured that the cross-sectional sample autocorrelation functions that were used to identify model structures employed an identical time period for each firm.

Table 1
Sample Selection Criteria

Firms listed on the COMPUSTAT industrial data base for which the following data were available:

COMPUSTAT Quarterly	*1985 Sample*	*1986 Sample*
Operating income before depreciation		
Interest expense or long-term debt		
Tax expense		
Deferred taxes		
Current assets		
Cash		
Inventory		
Receivables		
Current liabilities		
Short-term debt		
COMPUSTAT Annual Interest Expense[a]	165	123
December 31 Year End	109	80
Reduced Samples[b]	80	66

Note: The 1985 sample provided quarterly observations from 1976 through 1985. The 1986 sample provided observations from 1976 through 1986. In both samples, the last four quarters were withheld to assess predictive ability.

[a] Because quarterly interest expense was unavailable for a relatively large number of firms, annual interest expense was allocated to quarters in proportion to quarterly long-term debt when necessary.

[b] Our primary test samples were reduced further because of data requirements for the additional independent variables (contemporaneous values of sales and earnings; lagged values of sales, cash from operations, current and noncurrent accruals; and capital expenditures) in the pooled, cross-sectional regression model developed by Wilson (1986, 1987).

Stober (1989) as follows: (1) WCF = operating income before depreciation minus interest expense and the current portion of income tax expense; (2) *CF* = working capital from operations minus increases in net working capital other than cash, and securities, net of short-term debt. As observed by Bernard and Stober (1989), these measures have the advantage of being consistently calculated across firms.

II. Descriptive Evidence

In panel A of table 2, we present the cross-sectional sample autocorrelation (SACF) and partial autocorrelation functions (PACF) for the levels of the *CF* and WCF series. These functions were computed over 36 quarters, beginning with the first quarter of 1976 and ending with the fourth quarter of 1984, for each of the 109 firms in our 1985 sample.[4]

The SACFs for the *CF* series exhibit virtually no autocorrelation at the first three adjacent lags and only seasonal autocorrelation at lags 4, 8, and 12. In contrast, the WCF series exhibit *both* adjacent and seasonal lag patterns. One inference consistent

---

[4] We have withheld the eight quarters in 1985 and 1986 for predictive purposes.

Table 2
Cross-Sectional Sample Autocorrelations and Partial Autocorrelations; 1976–84
(means and standard deviations)

Panel A. Levels of Series ($d=0$, $D=0$):

	Lags											
	1	2	3	4	5	6	7	8	9	10	11	12
CF ($n=109$)												
SACF	−.006	.001	−.016	.307	−.040	−.072	−.032	.220	−.057	−.059	−.068	.138
	(.260)	(.255)	(.228)	(.253)	(.190)	(.219)	(.176)	(.228)	(.163)	(.184)	(.151)	(.196)
PACF	−.006	−.065	−.025	.221	−.073	−.096	−.014	.048	−.034	−.029	−.039	−.015
	(.260)	(.226)	(.243)	(.217)	(.178)	(.144)	(.144)	(.144)	(.152)	(.129)	(.125)	(.124)
WCF ($n=109$)												
SACF	.454	.322	.257	.343	.116	.051	.012	.113	−.019	−.055	−.060	.027
	(.280)	(.327)	(.245)	(.230)	(.228)	(.248)	(.197)	(.232)	(.152)	(.177)	(.136)	(.190)
PACF	.454	.067	.026	.137	−.145	−.068	−.028	.024	−.055	−.039	−.030	−.037
	(.280)	(.245)	(.183)	(.230)	(.156)	(.148)	(.118)	(.138)	(.121)	(.127)	(.125)	(.112)

Continued

Table 2—*Continued*
Cross-Sectional Sample Autocorrelations and Partial Autocorrelations; 1976–84
(means and standard deviations)

Panel B. Seasonal Differences of Series (*d*=0, *D*=1):

	Lags											
	1	2	3	4	5	6	7	8	9	10	11	12
CF (*n*=109)												
SACF	.043	.040	−.018	−.359	−.040	−.041	.008	−.006	.008	.031	−.012	−.036
	(.259)	(.139)	(.179)	(.178)	(.197)	(.170)	(.165)	(.182)	(.185)	(.156)	(.172)	(.152)
PACF	.043	−.035	−.037	−.365	−.015	−.071	.002	−.187	−.009	−.019	−.016	−.128
	(.259)	(.151)	(.177)	(.157)	(.193)	(.132)	(.150)	(.145)	(.151)	(.120)	(.142)	(.121)
WCF (*n*=109)												
SACF	.361	.188	.055	−.253	−.102	−.076	−.092	−.081	−.041	−.040	−.025	−.027
	(.246)	(.197)	(.179)	(.204)	(.179)	(.183)	(.169)	(.183)	(.164)	(.160)	(.157)	(.146)
PACF	.361	−.014	−.066	−.319	.092	−.021	−.046	−.161	.058	−.047	−.032	−.113
	(.246)	(.161)	(.168)	(.187)	(.159)	(.151)	(.137)	(.147)	(.136)	(.124)	(.121)	(.128)

Note: *d*=consecutive differencing,
D=seasonal differencing,
SACF=autocorrelation function,
PACF=partial autocorrelation function,
CF=cash flow per share series, and
WCF=working capital flow per share series.

with such time-series behavior is that short-term accruals (inventory, receivables, and payables) included in the WCF calculations may induce the adjacent lag patterns through smoothing characteristics. In fact, Beaver (1970) discusses how many aspects of accounting measurement may induce autocorrelation into such series even if the underlying process is serially independent. We reiterate that such behavior is noticeably absent from the CF series, which is free from such accruals.

Panel B presents similar information for the seasonally differenced ($d=0$, $D=1$) series.[5] The *CF* series, in marked contrast to the WCF series, exhibits no autocorrelation at the nonseasonal lags, but a singular spike at lag 4 of the SACF underscores the purely seasonal nature of the *CF* series, given the lack of autocorrelation at the first three lags. In comparison, the seasonally differenced WCF series depicts spikes at the first two adjacent lags as well as at the seasonal lag 4. In general, this pattern is consistent with the dual process characterization for quarterly earnings data popularized by Foster (1977) and Griffin (1977) and underscores the similarity of the WCF series to quarterly earnings.

The SACF patterns displayed in table 2 led to the identification of the following models for the *CF* series:[6] (1) (000)×(100), a seasonal autoregressive model (SAR); (2) (000)×(011), a seasonally differenced, seasonal moving average process (SMA); and (3) (000)×(010), a seasonal random walk model (SRW). The residual series for the SAR and SMA models did not contain significant autocorrelation. However, the residuals for the SRW model did, and it was therefore, dropped from further consideration.[7]

We also estimated parameters for the following models for the WCF series: (1) (100)×(100), an autoregressive and seasonal autoregressive ARIMA model (ASA) and (2) (011)×(011), the Griffin-Watts (GW) characterization.[8] The residual series for both models did not contain significant autocorrelation.[9] In the next section we assess the predictive ability of these models in a holdout period not used for model identification.

III. Predictive Ability

In this section we present results on the predictive ability of the *CF* and WCF series for our sample firms. For the *CF* series, we assess the performance of ARIMA-based models versus the pooled, cross-sectional regression model developed by Wilson (1987) and examined by Bernard and Stober (1989). This is subsequently referred to as the *W*-model. For comparative purposes, we also tested the predictive ability of two alternative time-series, multiple regression models (*TS*1 and *TS*2) that, like the *W*-model, incorporate recent earnings (*NI*) numbers in the expectation models for the *CF* series.

---

[5] The SACF and PACF of the consecutively differenced ($d=1$, $D=0$) as well as consecutively *and* seasonally differenced ($d=1$, $D=1$) series did not provide additional insights and are not presented.

[6] We adopt the familiar ARIMA notation wherein p, P indicate regular and seasonal autoregressive parameters, d, D indicate consecutive and seasonal differencing, and q, Q indicate regular and seasonal moving average parameters.

[7] Specifically, the critical value of the Q-statistic for the SAR and SMA models is 19.67 ($\alpha=0.05$, $dF=11$), with mean Q-statistics of 14.72 for SAR and 14.01 for SMA. The critical value of the Q-statistic for the SRW model is 21.03 ($\alpha=0.05$, $dF=12$), with a mean Q-statistic of 21.83.

[8] We also tested the (100)×(011) Brown-Rozeff model on the WCF series, but it proved inferior to both the GW and ASA models and was dropped from futher consideration.

[9] Specifically, the critical value of the Q-statistic for the ASA and GW models is 18.31 ($\alpha=0.05$, $dF=10$), with mean Q-statistics of 10.02 for ASA and 10.45 for GW.

The structural forms of these models are as follows (with variables scaled on a per-share basis):

$$(TS1)\quad CF_t = \Theta(CF_{t-1} - CF_{t-5}) + NI_t + u_t,$$

$$(TS2)\quad CF_t = \Theta(CF_{t-1} - CF_{t-5}) + \phi(NI_t - NI_{t-4}) + u_t,$$

where:

Θ = the firm-specific cash flow parameter,
ϕ = the firm-specific earnings parameter, and
u_t = the current disturbance term.

Finally, the GW and ASA ARIMA models are compared to the cross-sectional regression models in the prediction of WCF.

One-step-ahead quarterly *CF* and WCF predictions were generated in an *ex ante* fashion for all expectation models discussed above. Models were estimated on data beginning with the first quarter of 1976 and ending with the fourth quarter of 1984. The eight quarters in the 1985–86 interval were used as the prediction period. Parameters were estimated originally over the 1976–84 period to generate predictions for the first quarter of 1985. We then added the most recent observation and sequentially reestimated the parameters prior to forming each additional one-step-ahead prediction over the 1985–86 period. We computed two error metrics, mean absolute percentage error (MAPE) and mean squared error.[10] Since our statistical tests were insensitive to the choice of error metric, only the MAPE is reported here.

The predictive ability test involved operationalizing the *W*-model, for the CF and WCF series. Wilson (1987, 303) details the independent variables employed in this expectation model. Briefly, the first set of independent variables includes contemporaneous values of two variables announced in quarter *t*: sales and net earnings. The second set includes combinations of lagged values reported in two previous quarters ($t-1$ and $t-4$) and the sum of values for quarters $t-2$ and $t-3$ for the following variables: sales, cash from operations, current accruals, and noncurrent accruals. The last independent variable employed was the value of capital expenditures for the most recent annual period. All variables were scaled by total assets in a way similar to that of Wilson (1987) and Bernard and Stober (1989).

In these prior studies, data were pooled cross-sectionally across the entire data base in the estimation of the parameters for the independent variables in the *W*-model. An advantage to this procedure, which assumes intertemporal stability of the *W*-model in its parameters, is reduction in sampling variation in parameter estimation. However, this process is not a purely predictive one because the cross-sectionally derived parameters are estimated with full knowledge of the stipulated values of the independent variables across the prediction period. This precludes operationalizing the *W*-model in a true *ex ante* fashion.

To assess whether predictive ability is sensitive to alternative estimation procedures, we estimated the *W*-model parameters in two ways. First, we employed the procedure discussed above, allowing all data, including that for the prediction period 1985–86, to be used for parameter estimation. We refer to this as the in-sample *W*-model (WI). Second, we ran the model in a true *ex ante* fashion without allowing future values of the independent variables in the pooled cross-section. This is labeled the out-of-sample *W*-model (WO). The additional data requirements for the host of independent variables needed in these models reduced the number of firms to 80 (1985 sample) and 66 (1986 sample).

---

[10] These error metrics have been employed extensively in the forecasting literature; see, e.g., Bathke et al. (1989).

Table 3 presents the predictive results for the CF and the WCF series. All forecast errors greater than 100 percent were truncated to 100 percent. Specifically, the table presents truncated MAPE error metrics across models, series, and years in the holdout period.[11]

Panel A of table 3 shows the predictive performance of the regression models and the time-series models on the CF series. The best models, on the basis of the pooled MAPE metric, were the SAR (0.599) and SMA (0.621) ARIMA models; the models with the greatest errors were WO (0.715) and WI (0.699). More specifically, the SAR model's dominance over both *W*-models is evident in six of eight quarters in the prediction period. The SMA model also displays considerable dominance over both *W*-models, and both SAR and SMA outperform *TS*1 and *TS*2, although these results are not quite as strong.

Evidently, the cross-firm constraints imposed on the parameters in Wilson's expectation model may be too restrictive when compared to the firm-specific parameter estimation employed in the SAR ARIMA model. Alternatively, the structural form of the cross-sectional model may be misspecified.

Panel A also reveals very similar predictive performance by the WI and WO models. Panel B indicates that there was only one significant difference in the eight quarterly comparisons; WI outperformed WO in the first quarter of 1985 ($\alpha = 0.05$). This result suggests that the CF predictions of the *W*-model are not sensitive to the alternative estimation procedures discussed earlier.

Panel C presents comparisons of the ASA and GW ARIMA models vis-à-vis WI and WO on the WCF series. Panel D reveals that, with the exception of the fourth quarter in 1986, both the GW and ASA models are statistically superior ($\alpha = 0.05$) to WI and WO for all remaining quarters. In contrast to the results for the CF series, WI systematically outperforms WO, which suggests that alternative estimation procedures have a significant impact on the WCF predictions.[12]

IV. Security Return Effects

The time-series models provide an alternative set of expectations for assessing seurity price effects of cash flows and working capital from operations. For each of the eight quarters of 1985 and 1986 we regressed unexpected funds flow measures from each of the models on abnormal security returns over a nine-day period centered on the release of firms' financial statements.[13] Consistent with previous research, unexpected

---

[11] In all cases reported in table 3, the time-series models for both series were truncated less frequently than alternative models. For example, the SAR model resulted in the smallest percentage of truncations for the CF series (30.3%), while the WO model had the greatest (43.8%). Likewise, the ASA model exhibited the smallest percentage of truncations for the WCF series (14.5%) while the WO model had the greatest (37.8%).

[12] Similar results are obtained when prediction errors are scaled by total assets.

[13] Financial statement release dates were the earlier of the 10-K or annual report date for fourth quarter reports. For quarters 1 through 3, the 10-Q date was used. All dates were determined by reference to the date the SEC received the report. Also, further reductions in sample size were caused by CRSP data availability and the removal of firms that publicly disclosed pre-report financial data. We are grateful to Vic Bernard for providing this latter information.

Table 3
MAPEs of One-Step-Ahead Predictions

Panel A. CF Series:

		1985 Sample (n=80)				1986 Sample (n=66)				
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Pooled
SAR	(000)×(100)	.653	.550	.553	.556	.743	.577	.569	.610	.599
SMA	(000)×(011)	.716	.611	.570	.552	.732	.615	.557	.631	.621
WI		.774	.691	.680	.649	.743	.699	.685	.671	.699
WO		.858	.710	.644	.683	.807	.702	.664	.643	.715
TS1		.697	.575	.576	.682	.763	.598	.539	.626	.632
TS2		.682	.614	.653	.673	.789	.659	.654	.676	.673

Panel B. Wilcoxon Matched Pair Comparisons:

(First model results in significantly smaller MAPE, $\alpha=0.05$, than second model.)

1985 Sample (n=80)				1986 Sample (n=66)			
Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
SAR-SMA SAR-WI SAR-WO SMA-WI SMA-WO WI-WO TS1-WI TS1-WO TS2-WI TS2-WO	SAR-WI SAR-WO SMA-WI SMA-WO TS1-WI TS1-WO TS1-TS2 TS2-WI TS2-WO	SAR-WI SAR-TS2 SMA-WI TS1-WI TS1-TS2	SAR-WI SAR-WO SAR-TS1 SAR-TS2 SMA-WI SMA-WO SMA-TS1 SMA-TS2	SAR-WO SMA-WI SMA-WO	SAR-WI SAR-WO SMA-WI TS1-WI TS1-WO TS1-TS2 TS2-WI	SAR-WI SAR-WO SAR-TS2 SMA-WI SMA-WO TS1-WI TS1-WO TS1-TS2 TS2-WI	None

Continued

Table 3—*Continued*
MAPEs of One-Step-Ahead Predictions

Panel C. WCF Series:

		1985 Sample (*n*=80)				*1986 Sample* (*n*=66)				
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Pooled
GW	(011)×(011)	.396	.320	.343	.424	.496	.288	.345	.546	.393
ASA	(100)×(100)	.357	.332	.367	.467	.468	.289	.334	.474	.386
WI		.619	.549	.555	.593	.678	.573	.559	.599	.590
WO		.657	.724	.594	.616	.701	.651	.609	.627	.647

Panel D. Wilcoxon Matched Pair Comparisons:

(First model results in significantly smaller MAPE, $\alpha=0.05$, than second model.)

1985 Sample (*n*=80)				*1986 Sample* (*n*=66)			
Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
GW-WI GW-WO ASA-WI ASA-WO	GW-WI GW-WO ASA-WI ASA-WO WI-WO	GW-WI GW-WO GW-ASA ASA-WI ASA-WO	GW-WI GW-WO ASA-WI ASA-WO	GW-WI GW-WO ASA-WI ASA-WO	GW-WI GW-WO ASA-WI ASA-WO WI-WO	GW-WI GW-WO ASA-WI ASA-WO WI-WO	GW-WI ASA-GW ASA-WI ASA-WO WI-WO

Note: SAR=seasonal autoregressive model,
SMA=seasonally differenced, seasonal moving average model,
ASA=autoregressive, seasonal autoregressive model,
GW=Griffin-Watts model,
TS1=time-series multiple regression model 1,
TS2=time-series multiple regression model 2,
WI=Wilson-in-sample model, and
WO=Wilson-out-of-sample model.

WCF failed to provide any significant explanatory power after adjusting for outliers. This lack of information content is consistent with the similarity of the time-series properties of earnings and WCF, and the relative predictability of long-term accruals.

Despite the fact that the SAR and SMA time-series models for cash flows outperformed the cross-sectional models with respect to predictive ability, this result did not translate into consistent associations with the security return metric. As in previous research, significant effects for unexpected CF were discovered in only some of the quarters tested. Moreover, when data were pooled across all eight quarters, the overall explanatory power of all the models presented in this study was uniformly low. Although these results suggest that Wilson's (1987) finding on the information content of cash flow may be sensitive to the expectation model employed, it was unclear from our study whether the difference might also be attributed to sampling variation, the time period analyzed, or other unspecified factors.

V. Conclusions

We demonstrate that univariate time-series models of CF and WCF generate more accurate forecasts than the multivariate cross-sectional regression models used in prior research (see, e.g., Bernard and Stober 1989; Wilson 1986, 1987). Specifically, the SAR ARIMA model is a candidate for a parsimonious prediction model for the CF series, and both the ASA and GW ARIMA models perform well on the WCF series. This outcome may appear to be somewhat counterintuitive initially given the multitude of earnings and accrual data employed in the multivariate models. However, a plausible explanation for the domination of the univariate models pertains to the cross-sectional restriction on the coefficients in the regression models. Firm-specific estimation of the coefficients in the ARIMA models is a likely explanation for the improvement in predictive ability reported here. We do not assert, however, that the relatively parsimonious ARIMA models identified are necessarily optimal for funds flow variables. The cross-sectional restriction on model coefficients as well as potential model misspecification may have prevented the multivariate models from capturing the information content of accrual accounting variables. Additionally, the analysis we conducted was restricted to firms with a December 31 fiscal year end. Therefore, the results may not be generalizable to firms that employ different year ends.

Our descriptive analyses provide a necessary first step in the examination of the time-series properties of quarterly CF and WCF data. It appears that purely seasonal ARIMA models were able to capture the autocorrelation patterns in the CF series. In sharp contrast, the ARIMA models identified for the WCF series were virtually identical to the well-documented premier ARIMA models for quarterly earnings. This finding suggests that it may be appropriate to model the CF and WCF series with models having distinct structural forms.

Additional research efforts are needed to identify multivariate cash flow models that overcome the inherent limitations of the cross-sectional regression models popularized in the literature. Such models may be able to exploit the likely potential information in accounts receivable, accounts payable, inventory, and other accruals in forecasting cash flow. Finally, the development of industry-specific models would be particularly interesting.

References

Bathke, A. W., K. S. Lorek, and G. L. Willinger. 1989. Firm-size and the predictive ability of quarterly earnings data. *The Accounting Review* 64 (January): 49–68.

Beaver, W. H. 1970. The time series behavior of earnings. *Journal of Accounting Research* 8 (Supplement): 62–99.

Bernard, V., and T. L. Stober. 1989. The timing, amount, and nature of information reflected in cash flows and accruals. *The Accounting Review* 64 (October): 624–52.

Bowen, R., D. Burgstahler, and L. Daley. 1986. Evidence on the relationships between earnings and various measures of cash flow. *The Accounting Review* 61 (October): 713–25.

——, ——, and ——. 1987. The incremental information of accrual income versus cash flows. *The Accounting Review* 62 (October): 723–47.

Financial Accounting Standards Board. 1978. *Statement of Financial Accounting Concepts No. 1:* Stamford, CT: FASB.

Foster, G. 1977. Quarterly accounting data: Time-series properties and predictive-ability results. *The Accounting Review* 52 (January): 1–21.

Griffin, P. A. 1977. The time-series behavior of quarterly earnings: Preliminary evidence. *Journal of Accounting Research* 15 (Spring): 71–83.

Schaefer, T. F., and M. D. Kennelley. 1986. Alternative cash flow measures and risk-adjusted returns. *Journal of Accounting, Auditing & Finance* (October): 278–87.

Wilson, G. P. 1986. The relative information content of accruals and cash flows: Combined evidence at the earnings announcement and annual report release date. *Journal of Accounting Research* 24 (Supplement): 165–200.

——. 1987. The incremental information content of the accrual and funds components of earnings after controlling for earnings. *The Accounting Review* 62 (April): 293–322.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 164–175

The Influence of the Source of Professional Standards on Juror Decision Making

Frank A. Buckless
Robert L. Peace
North Carolina State University

SYNOPSIS AND INTRODUCTION: Auditor liability is a growing concern of public accounting firms (Collins 1985; Flynn et al. 1990; Mednick 1987; Minow 1984). In lawsuits brought by clients for breach of contract and tort actions for negligent execution of an audit, as well as in an increasing number of lawsuits brought by third parties, a common defense is the auditor's compliance with generally accepted auditing standards (GAAS). In this study, the source of professional auditing standards was manipulated to examine its influence on jurors' decisions.

Four fact patterns were presented to members of jury pools called to jury duty. Prospective jurors responded to a fact pattern by voting in favor of either the plaintiff/client or the defendant/auditor. The fact patterns were manipulated to compare jurors' responses to standards of auditing performance established by the federal government versus those established by the auditing profession. In addition, a comparison was made of jurors' responses to two sets of instructions provided by the judge: either that standards alone set the required level of performance or that standards are only a part of the evidence to be considered in establishing the required level of performance. At trial, a judge's instructions to the jury before its deliberations can be influenced by counsel for the public accounting firm as to the evidentiary weight assigned to professional standards.

The results that follow indicate that jurors are more likely to rule against a CPA firm when the profession's standards are offered in defense and less likely to rule against it when government standards of performance are offered in defense. Also, as expected, jurors are more likely to accept standards in defense, whether established by the government or the pro-

---

The first author gratefully acknowledges the funding provided by the KPMG Peat Marwick Foundation Research Fellowship Program.

Submitted November 1991.
Accepted September 1992.

fession, if the judge identifies the standards as the only criterion for an expected level of performance rather than as only part of the evidence.

Key Words: *Auditing standards, Auditor liability, Jury perceptions, Negligence.*

Data Availability: *Data are available from the first author upon request.*

I. Background

A PRIMARY defense used by auditing firms sued for negligence in the conduct of an audit is compliance with professional auditing standards. The auditing profession, through the American Institute of Certified Public Accountants AICPA), sets professional auditing standards of performance. Currently, the AICPA as designated the Auditing Standards Board (ASB) as the senior technical body harged to issue pronouncements on auditing standards. Whether compliance with the tandards issued by the ASB is an adequate defense against malpractice lawsuits epends, in part, on the public's confidence in the auditing profession's standard-set- ing process. As Jennings et al. (1991) note, auditor liability may depend most impor- antly on the attitudes of jurors and judges toward the accounting profession.

Auditor success rates, when a lawsuit goes to trial, have been lower than those of ther professions (Palmrose 1991). The nature of the standard-setting organization and rocedures may give the appearance that the ASB favors the needs of auditors over hose of audit clients and other interested parties (Kripke 1988).[1] The primary purpose f this study is to examine whether the source of auditing standards affects jurors' deci- ions in negligence cases in which compliance with professional standards has been roved.[2]

Prior to deliberating on a verdict, jurors are instructed by the trial judge. The nited States Supreme Court has noted that our jury trial system depends on the ‘crucial assumption . . . that juries will follow the instructions given them by the trial udge" (Tanford 1990, 73). A review of appellate court cases leads to an uncertain con- lusion regarding the probative value of GAAS assigned by the courts (Jacobson v. Peat, arwick, Mitchell, & Co., 445 F. Supp. 518, 1977; Securities and Exchange Commis- ion v. Arthur Young & Co., 590 F. 2d 785, 1979; Quantum Overseas, N.V. v. Touche oss & Co., 663 F. Supp. 658, SDNY 1987; Maduff Mortgage v. Deloitte Haskins & Sells, 79 P. 2d 1083, 1989). A secondary purpose of the study will be to examine how a udge's instructions regarding compliance with professional standards affects jurors' ecisions on negligence when compliance with professional standards has been roved.

---

[1] The Treadway Commission's (National Commission on Fraudulent Financial Reporting 1987) recom- endation of a more diverse ASB membership recognizes the possibility that the public might perceive a con- lict of interest when auditing professionals prescribe their own standards.

[2] The Seventh Amendment to the United States Constitution provides, in part, that "in suits at common law here the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved . . ." Austin 1984). Furthermore, decisions on negligence lawsuits against auditors are normally rendered by a jury Palmrose 1991).

II. Research Hypotheses

Source of Professional Standards

Previous research in the social sciences suggests that jurors will not use laws or criteria that they believe are unfair or unjust in making decisions (Kassin and Wrightsman 1988; Myers 1979; Mysliwiec 1974; Pennington and Hastie 1990). Greenberg (1975) suggests that the appearance of fairness may be as important as actual fairness. Leventhal (1980) developed a model of how decision makers' perceptions of fairness, and thus their decisions, can be influenced. In this model, decision makers' perceptions of fairness are affected by both the criteria used to allocate punishment or reward and by the procedures followed to establish the decision criteria (Leventhal 1980; Lind and Tyler 1988). Compliance with GAAS is the decision criterion auditors assert in defense of their performance. Standard setting by the ASB (self-regulation) is the procedure used to establish the decision criteria (GAAS).

The Leventhal model suggests that when a jury evaluates the appropriateness or fairness of professional standards, it will also evaluate the procedure used to establish the performance standards (Leventhal 1980; Lind and Tyler 1988). Leventhal (1980) identifies six elements of procedural fairness that could influence individual perceptions: bias suppression, representativeness, consistency, accuracy, correctability, and ethicality.

Of these six factors, two with particular relevance to the standard-setting process employed by the auditing profession are bias suppression and representativeness. Auditing standards established by professionals, whose performance is judged by adherence to these same standards, may give the impression of bias in favor of the needs of auditors rather than the needs of audit clients and the public. As Leventhal (1980, 41) notes, "An individual is likely to believe that procedural fairness is violated when there is unrestrained self-interest...." Representativeness demands that the regulatory process incorporate the basic concerns, values, and outlook of important subgroups in the affected population. This model suggests that, for jurors to perceive the regulatory process as fair, the regulating organization should include important subgroup participants or disinterested outside participants. Tyler (1989) conducted an empirical study of the effect of bias suppression and representativeness (referred to as neutrality and control) on perceptions of fairness, and found that both these factors influence an individual's perception of fairness.

A related line of research examines the influence of the credibility of the information provider (Schum 1975, 1977; Schum and DuCharme 1971; Schum and Martin 1982; Schum and Pfeiffer 1973) and also suggests that the source of the professional standards is likely to influence jurors' decisions. According to the source credibility paradigm developed by Schum and associates, as jurors' evaluations of the credibility of a source of information decreases, the influence of that information on jurors' decisions also decreases. Empirical research also supports the suggestion that an unbiased source of evidence, one with no stake in the outcome, is more persuasive than a biased source (Reinard 1988).

The research hypothesis related to the composition of the regulating organization is as follows:[3]

---

[3] Government and professional standards were used as the two forms of regulation in this study since the current regulatory environment is a mixture of the two. Additionally, these two forms of regulation are common and thus should minimize juror misunderstandings.

H1: Jurors are more likely to find in favor of a defendant/auditor in a negligence lawsuit when the auditing standards offered in defense are established by the federal government as opposed to the auditing profession.

Judge's Instructions

Blanck et al. (1990) found that a judge's expectation for a trial outcome can lead the jury to confirm this expectation. Other research indicates that instructions may be discounted or overlooked by jurors. It has been shown that jurors who appear to fully understand a judge's instructions will not follow the instructions if they find them unfair (Kalven and Zeisel 1966; Loftus 1987; Mysliwiec 1974).

The Leventhal (1980) model implies that the judge's instructions can influence jurors' decisions. In particular, the model suggests that a decision criterion supported by a legitimate outside authority can reduce a decision maker's concern for procedural fairness (i.e., bias suppression, representativeness, and other elements). Thus, the influence of the source of standards on jurors' decisions may be reduced when the judge informs the jurors that the professional standards themselves are the appropriate (or only) consideration.

The research hypothesis related to the judge's instructions is as follows.

H2: Jurors are more likely to find in favor of the defendant/auditor when the judge instructs the jurors that the auditing standards are the only standard that should be used to establish the CPA's required level of performance, as opposed to when the judge instructs the jurors that the auditing standards are only one element of the evidence that should be considered.

III. Research Method

Participants

Jurors who were part of a county jury pool for the period July 8 through August 16, 1991, were asked to participate in the study.[4] This is the same jury pool that would be used in an actual auditor negligence lawsuit in this county. A total of 300 case instruments were handed out at the courthouse to county residents who were summoned for jury duty. Written responses were obtained for four different case instruments describing circumstances in a negligence lawsuit brought against a public accounting firm. Participants were asked to decide for the plaintiff or the public accounting firm on a ballot accompanying each case instrument. Two hundred seventy-five useable responses were collected and included in the study's data. Demographic information from a questionnaire that accompanied each case instrument is provided in table 1.[5]

Case Instruments and Experimental Design

Each participating juror was asked to read one of four versions of a one-page hypothetical lawsuit brought by a client alleging negligence against a public accounting

---

[4] Residents in the county are called from a jury list to serve in a pool or venire of jurors, from which individual jurors are called for specific trials. The source of names used in the county jury system is the Board of Elections master file of registered voters and a file of county residents from the State Department of Motor Vehicles. People under 18 years of age are dropped from the list, as are those with no first or middle name and those individuals who were previously exempt from jury service. This random file is sorted into numerical order, and each jury venire is then selected sequentially.

[5] The demographic data are consistent with the highly educated and professionally employed work force in the county.

Table 1
Demographic Information on Participants

Panel A. Participants' Responses to Quantitative Questions:

	Mean	*S.D.*	*Min.*	*Max.*
Age	40	12.3	18	73
Family income	$55,457	$56,835	$0	$760,000
Years of education	14.6	2.7	2	25

Panel B. Participants' Responses to Qualitative Questions:

Prior civil jury trial experience of participants:	
Never served on civil jury trial	83.0%
Previously served on civil jury trial	17.0%
	100.0%
Personal investment activity of participants:	
Never invest in stocks/bonds	45.6%
Rarely invest in stocks/bonds	15.3%
Sometimes invest in stocks/bonds	24.5%
Invest in stocks/bonds most of the time	6.2%
Invest in stocks/bonds all of the time	8.4%
	100.0%
Participants who were:	
Female	52.6%
Male	47.4%
	100.0%
Participants who were:	
Married	64.3%
Divorced/separated	9.6%
Single	26.1%
	100.0%
Participants with:	
Conservative viewpoint	53.7%
Liberal viewpoint	46.3%
	100.0%
Participants who were:	
Housewives	5.6%
Retired	4.1%
Employed	90.3%
	100.0%

firm. Each version is identical on the facts of the audit and the negligence claim.[6] Jurors were asked to reach a decision for either the plaintiff/client or the defendant/auditor.

The research design represented a 2 × 2 fixed-effect factorial design with two different sources of standards and two different instructions provided by the judge. In two cases, the defense asserted that the public accounting firm complied with GAAS. In the other two cases, the defense asserted compliance with government auditing standards. The judge's instruction was either that compliance with standards is the only criterion to use in evaluating the public accounting firm's performance or compliance is to be considered along with any other evidence as a measure of performance.[7]

The court jury clerk agreed to distribute the case instruments to the jurors, and there was no direct communication between the researchers and the juror participants. The four scenarios were randomly ordered prior to their delivery to the jury clerk, who read the written instructions to respondents and collected the completed instruments.

IV. Results

Effect of Source of Standards and Judge's Instructions

A maximum likelihood logistic regression was used to investigate the influence of the source of standards and the judge's instructions on the jurors' decisions. The dependent variable for this analysis was the jurors' decisions. The independent variables were the two different sources of standards and two different instructions. Effect dummy coding (i.e., -1, 1) was used for the two categorical independent variables. A categorization of jurors' responses identified by the source of standards and by the judge's instructions is provided in table 2. The maximum likelihood logistic regression analysis is displayed in table 3, panel A. The R-value for this logistic regression model is 0.22.[8] The model chi-square test indicates that the overall logistic regression model is statistically significant ($p < 0.001$).

The results in tables 2 and 3 indicate statistically significant differences in the hypothesized direction for both independent variables. Jurors were more likely to find in favor of the defendant/auditor when the standards were established by the federal government ($p < 0.035$) as opposed to the auditing profession. Additionally, jurors were more likely to find in favor of the defendant/auditor when the judge instructed the jurors that compliance with standards is the only appropriate criterion for evaluating performance, as opposed to being only part of the evidence ($p < 0.001$).[9]

[6] Pilot tests with 91 undergraduate students at a large state university resulted in only minor modifications to the case instruments.

[7] Copies of the four versions of the case instrument can be obtained from the first author upon request.

[8] The R-value is a goodness-of-fit measure for the logistic regression model and is similar to the multiple correlation coefficient used in a linear regression analysis. The R-value represents the proportion of the log likelihood explained by the model (Atkinson 1980).

[9] A chi-square test was also used to further corroborate the results of the logistic regression. The test was performed by collapsing the cells for one independent variable and performing a chi-square test on the other independent variable. The results are consistent with the results of the logistic regression (see notes to table 3).

To evaluate the possibility that the results were driven by juror familiarity with the auditing profession, we conducted a separate analysis that excludes those respondents who could be reasonably expected to have knowledge of the auditing profession, on the basis of their responses to demographic questions (i.e., accountants, those providing financial services, business executives, and those investing in the stock market most or all of the time). The results with this reduced sample were consistent with the results reported.

Table 2
Contingency Tables of Juror Responses

Panel A. Juror Decisions by Source of Standards and Judge's Instruction:

	Profession		*Government*		
	Part of Evidence	*Only Evidence*	*Part of Evidence*	*Only Evidence*	
For CPA	36 (55%)	51 (74%)	43 (61%)	62 (87%)	192 (70%)
For Plaintiff	29 (45%)	18 (26%)	27 (39%)	9 (13%)	83 (30%)
	65 (100%)	69 (100%)	70 (100%)	71 (100%)	275 (100%)

Panel B. Juror Decisions by Source of Standards:

	Profession	*Government*	
For CPA	87 (65%)	105 (74%)	192 (70%)
For Plaintiff	47 (35%)	36 (26%)	83 (30%)
	134 (100%)	141 (100%)	275 (100%)

$\chi^2=2.969$, df=1, $p<0.043$, one-tailed.

Panel C. Juror Decisions by Judge's Instruction:

	Part of Evidence	*Only Evidence*	
For CPA	79 (59%)	113 (81%)	192 (70%)
For Plaintiff	56 (41%)	27 (19%)	83 (30%)
	135 (100%)	140 (100%)	275 (100%)

$\chi^2=16.068$, df=1, $p<0.001$, one-tailed.

Effect of Attitudinal and Demographic Variables

Prior research does not provide sufficient information for *a priori* predictions of the relationship of attitudinal and demographic variables with jurors' decisions (Boyll 1991; Kaplan and Lynn 1978; Kassin and Wrightsman 1988; Saks 1987; Vinson 1986). Demographic information has been found to interact with the specifics of a case in complex ways, which makes prediction of jurors' decisions from this information extremely difficult (Boyll 1991). In a data base consisting of more than 14,000 juror

Table 3
Maximum Likelihood Logistic Regression Summary Tables

Panel A. Jurors' Verdict Regressed on Source of Standards and Judge's Instruction (n=275):

Effect	*df*	*Coef.*	*Chi-square*	*P<*
Intercept	1	0.89	41.11	0.001
Source of standards (PS=−1, GS=1)	1	0.25	3.30	0.035*
Judge's instruction (PE=−1, OE=1)	1	0.55	15.78	0.001*
Model	2	n/a	19.64	0.001

Panel B. Jurors' Verdict Regressed on Source of Standards, Judge's Instruction, and Demographic Variables (n=207):

Effect	*df*	*Coef.*	*Chi-square*	*P<*
Intercept	1	0.72	0.41	0.520
Source of standards (PS=−1, GS=1)	1	0.37	4.08	0.022*
Judge's instruction (PE=−1, OE=1)	1	0.68	13.97	0.001*
Gender (M=−1, F=1)	1	0.06	0.10	0.750
Marital status (SM=−1, DM=0, MM=1)	1	0.41	1.83	0.176
Marital status (SM=−1, DM=1, MM=0)	1	0.24	0.33	0.564
Age	1	−0.06	12.09	0.001
Years of education	1	0.16	6.06	0.014
Family income	1	−0.00	0.00	0.992
Investment activity	1	0.11	0.54	0.463
Political viewpoint (L=−1, C=1)	1	−0.11	0.37	0.543
Civil jury experience	1	0.03	0.06	0.809
Model	11	n/a	39.62	0.001

Notes: An asterisk indicates a one-tailed test (two-tailed otherwise). The coding used for the categorical verdict variable was: plaintiff=0, defendant (CPA)=1.

PS=professional standards,
GS=government standards,
PE=part of evidence,
OE=only evidence,
M=male,
F=female,
SM=single,
DM=divorced/separated,
MM=married,
C=conservative, and
L=liberal.

interviews, demographics consistently failed to be statistically significant as a predictor of either verdict or dollar damages (Vinson 1986). Given the results of prior psychological and legal research, we conducted an exploratory analysis of the relationships among attitudinal and demographic variables and the jurors' decisions.

A maximum likelihood logistic regression was used to examine the relationship of the attitudinal and demographic variables with jurors' decisions. The dependent variables in this regression was the jurors' decisions. The independent variables were the source of standards and the judge's instructions, together with each juror's gender, marital status, age, years of education, family income, investments activity, and civil jury experience. The results of this analysis are presented in table 3, panel B.[10] The R-value for this logistic regression model is 0.27. The model chi-square test indicates that the overall logistic regression model is statistically significant ($p < 0.001$).

The results indicate statistically significant relationships for four of the ten variables in the analysis: source of standards, judge's instructions, each juror's age, and years of education.[11] The statistical significance of the source of standards and the judge's instruction is consistent with the previous logistic regression and supports the juror randomization method employed by the researchers.

The age variable was negatively associated with finding for the defendant/auditor, while the years of education were positively associated with finding for the defendant/auditor. The age variable indicates that the older the juror, the less likely he or she will find for the defendant/auditor. According to Boyll (1991), one of the more consistent findings regarding demographic information is that young jurors are more lenient in their verdicts. Our results are consistent with this finding.

The education variable indicates that better-educated jurors are more likely to find for the defendant/auditor. The existent psychology and legal literature does not provide an obvious explanation of this finding.[12]

Predictive Ability of Attitudinal and Demographic Variables

A second analysis was conducted to determine the usefulness of the manipulated and demographic variables in predicting jurors' decisions. The predictive capacity of the manipulated and demographic variables was tested by using a cross-validation procedure whereby each observation in the data set was classified with a logistic regression computed from the other observations in the data set (Geisser 1975). We present the predictive accuracy of the logistic regression model in table 4. A juror was predicted to find for the CPA if the estimated probability of this verdict was greater than 50 percent. The model predicts the juror's actual decisions with 70 percent accuracy (144/207).

The statistical significance of the model's classification accuracy was tested against a proportional-chance classification model, which classifies decisions by assuming random guessing based on known priors. Using the known frequencies of jurors' decisions, a proportional-chance model would correctly classify 60 percent of the decisions ($0.28^2 + 0.72^2$) in the sample. The statistical difference was tested with a T-statistic that approximates a chi-square distribution with one degree of freedom (Conover 1971; Mensah 1983). The T-statistic indicates a statistically significant difference between the logistic and proportional-chance models ($T = 4.23$, $p < 0.02$).

---

[10] The inclusion of the attitudinal and demographic variables reduced the number of participants who could be used in the analysis because of missing observations.

[11] Maximum likelihood logistic regressions also were run separately for each independent variable, with results that were consistent with the overall analysis.

[12] Prior research has indicated that better-educated individuals are less likely to sit as jurors in business cases (Hans 1989). As Ellsworth (1989) notes, jurors often are selected for incompetence. In our study, better-educated jurors were prone to find for the CPA. This would suggest that a randomly selected jury, which includes highly educated individuals, provides a low estimate of the number of jurors that would find for the plaintiff in an actual lawsuit.

Table 4
Logistical Regression Model Classification Accuracy

Panel A. All Independent Variables Included in the Logistic Regression Model:

		Actual		
		CPA	*Plaintiff*	
Predicted	CPA	130	44	174
	Plaintiff	19	14	33
		149	58	207

Decisions correct=70 percent.

Panel B. Only Significant Independent Variables Included in the Logistic Regression Model:

		Actual		
		CPA	*Plaintiff*	
Predicted	CPA	153	52	205
	Plaintiff	21	23	44
		174	75	249

Decisions correct=71 percent.

The final analysis examined predictive ability when only the source of standards, judge's instructions, jurors' ages, and jurors' years of education were included as independent variables in the logistic regression model. The previously discussed cross-validation procedure was also used to test the predictive capacity of the four independent variables. These results are also shown in table 4. The model predicts the jurors' actual decisions with 71 percent accuracy (176/249). Using the known frequencies of jurors' decisions, a proportional-chance model would correctly classify 58 percent of the decisions ($0.30^2+0.70^2$) in the sample. The *T*-statistic indicates a statistically significant difference between the logistic and proportional-chance models ($T=8.42$, $p<0.01$).

V. Discussion and Concluding Comments

The results indicate that jurors attached greater credibility to auditing standards established by the federal government than to those established by the auditing profession. Jurors' attitudes are affected when they perceive that a standard of performance is self-serving. These results are consistent with both the Leventhal model of justice judgment (Leventhal 1980; Lind and Tyler 1988) and empirical evidence (Tyler 1989). The results also emphasize the importance of a judge's instructions on the evidentiary value of compliance with auditing standards.

Although we employed a methodology commonly used in the social sciences

(Kassin and Wrightsman 1988; Pennington and Hastie 1990), it has limitations. The case instruments presented to the jurors contained less information than is normally presented at trial. Also, our study did not incorporate the *voir dire*[13] and group decision-making processes found in an actual jury deliberation.[14] Additionally, jurors' understanding of the auditing standard-setting process was measured indirectly, by using their response to demographic questions, rather than by a more rigorous direct manipulation or knowledge check. Finally, including a hybrid source of standards (i.e., a standard-setting organization consisting of both CPAs and non-CPAs) could have enhanced the implications of the study.

Our use of jurors as a source of data in an auditor litigation experiment raises a number of interesting questions for subsequent research. Is there an optimal mix of CPAs and non-CPAs that creates a credible standard-setting organization? Do jurors hold large and small CPA firms to the same standard of performance? Does the type of plaintiff (i.e., individual or corporation) affect jurors' verdicts? Does the plaintiff's previous relationship with the audit client affect jurors' verdicts? Does the performance of other services by an auditor for a client influence jurors' verdicts? Can auditors defend their performance when alternative accounting principles or practices are available to clients? What audit litigation factors are most critical in determining jurors' damage awards? For example, does the source of the standards influence the amount of punitive damages awarded by jurors? Can identifiable attitudinal and demographic characteristics be used to predict jurors' predispositions?

Future studies could increase experimental realism by including a greater range of facts and circumstances from actual lawsuits in the instrument. Mock trials based on actual cases could be enacted, videotaped, and presented to jurors (see, e.g., Pennington and Hastie 1990).

---

[13] *Voir dire* refers to the pre-trial juror selection process wherein each party exercises rights of peremptory challenge to jurors. In addition, counsel for each party challenges when he or she thinks that a cause for challenge is apparent, and the court then determines whether to sustain or overrule the challenge.

[14] Previous jury research has shown that, in cases similar to the one presented in the research instrument, in this paper, if a majority of jurors have an opinion before deliberation, that opinion will carry through to the jury's verdict (Kalven and Zeisel 1966; Pennington and Hastie 1990).

References

Atkinson, A. C. 1980. A note on the generalized information criterion for choice of a model. *Biometrika* 67 (August): 413–18.

Austin, A. D. 1984. *Complex Litigation Confronts the Jury System: A Case Study*. Frederick, MD: University Publications of America.

Blanck, P. D., R. Rosenthal, A. J. Hart, and F. Bernieri. 1990. The measure of the judge: An empirically based framework for exploring trial judges' behavior. *Iowa Law Review* 75 (March): 679–83.

Boyll, J. R. 1991. Psychological, cognitive, personality and interpersonal factors in jury verdicts. *Law and Psychology Review* 15 (Spring): 163–84.

Collins, S. H. 1985. Professional liability: The situation worsens. *Journal of Accountancy* 160 (November): 57–66.

Conover, W. J. 1971. *Practical Nonparametric Statistics*. New York: John Wiley & Sons.

Ellsworth, P. C. 1989. Are twelve heads better than one? *Law and Contemporary Problems* 52 (Autumn): 205–24

Flynn, J., S. B. Garland, C. Yang, and D. Foust. 1990. Pressure on professionals: From lawyers to accountants, privileges are being eroded. *Business Week* (July 23): 23–25.

Geisser, S. 1975. The predictive sample reuse method with applications. *Journal of the American Statistical Association* 70 (June): 320–28.

Greenberg, D. S. 1975. Peer review under fire. *Science Digest* 78 (September): 77–78.

Hans, V. P. 1989. The jury's response to business and corporate wrongdoing. *Law and Contemporary Problems* 52 (Autumn): 177–204.

Jennings, M. M., P. M. J. Reckers, and D. C. Kneer. 1991. The auditor's dilemma: The incongruous judicial notions of the auditing profession and actual auditor practice. *American Business Law Journal* 29 (Spring): 99–125.

Kalven, H., and H. Zeisel. 1966. *The American Jury*. Boston: Little, Brown.

Kaplan, M. F., and L. E. Lynn. 1978. Reducing the effects of juror bias. *Journal of Personality and Social Psychology* 36 (December): 1443–55.

Kassin, S. M., and L. S. Wrightsman. 1988. *The American Jury on Trial: Psychological Perspective*. New York: Hemisphere Publishing.

Kripke, H. 1988. Some reactions and some down-to-earth reflections on accountants' liability—A general discussion. *Journal of Accounting, Auditing & Finance* 3 (Fall): 359–67.

Leventhal, G. S. 1980. What should be done with equity theory? New approaches to the study of fairness in social relations. In *Social Exchange Theory*, edited by K. J. Gergen et al. New York: Plenum Press, 27–55.

Lind, E. A., and T. R. Tyler. 1988. *The Social Psychology of Procedural Justice*. New York: Plenum Press.

Loftus, E. 1987. Do jurors follow instructions?: An experiment using the randomized response technique to determine whether jurors obey the instructions not to discuss their case before deliberations. *University of Bridgeport Law Review* 8 (2): 359–64.

Mednick, R. 1987. Accountants' liability: Coping with the stampede to the courtroom. *Journal of Accountancy* 164 (September): 118–22.

Mensah, Y. M. 1983. The differential bankruptcy predictive ability of specific price level adjustments: Some empirical evidence. *The Accounting Review* 58 (April): 228–46.

Minow, N. N. 1984. Accountants' liability and the litigation explosion. *Journal of Accountancy* 158 (September): 70–86.

Myers, M. A. 1979. Rule departures and making law: Juries and their verdicts. *Law and Society Review* 13 (Spring): 781–97.

Mysliwiec, S. R. 1974. Toward principles of jury equity. *Yale Law Journal* 83 (April): 1025–54.

National Commission on Fraudulent Financial Reporting. 1987. Report of the National Commission on Fraudulent Financial Reporting. New York: AICPA.

Palmrose, Z. 1991. Trials of legal disputes involving independent auditors: Some empirical evidence. *Journal of Accounting Research* 29 (Supplement): 149–85.

Pennington, N., and R. Hastie. 1990. Practical implications of psychological research on jury decision making. *Personality and Social Psychology Bulletin* 16 (March): 90–105.

Reinard, J. C. 1988. The empirical study of the persuasive effects of evidence: The status after fifty years of research. *Human Communication Research* 15 (Fall): 3–59

Saks, M. J. 1987. Social scientists can't rig juries. In *In the Jury Book*, edited by L. S. Wrightsman et al. Newbury Park, CA: SAGE, 48–61.

Schum, D. A. 1975. The weighting of testimony in judicial proceedings from sources having reduced credibility. *Human Factors* 17 (2): 172–82.

———. 1977. Behavioral richness of cascaded inference models: Examples of jurisprudence. In *Cognitive Theory*, vol. 2, edited by N. J. Castellan et al. Hillsdale, NJ: Erlbaum, 149–73.

———. and W. M. DuCharme. 1971. Comments on the relationship between the impact and the reliability of evidence. *Organizational Behavior and Human Performance* 6 (March): 111–31.

———. and D. W. Martin. 1982. Formal and empirical research on cascaded inference in jurisprudence. *Law and Society Review* 17 (1): 105–51.

———. and P. E. Pfeiffer, 1973. Observer reliability and human inference. *IEEE Transactions on Reliability* 22 (August): 170–76.

Tanford, J. A. 1990. The law and psychology of jury instructions. *The Nebraska Law Review* 69 (1): 71–81.

Tyler, T. R. 1989. The psychology of procedural justice: A test of the group-value model. *Journal of Personality and Social Psychology* 57 (November): 830–38.

Vinson, D. E. 1986. *Jury Trials: The Psychology of Winning Strategy*. Charlottesville, VA: Michie Company.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 176–178

A Comment on "The Influence of the Source of Professional Standards on Juror Decision Making"

Jane F. Mutchler
Pennsylvania State University
Kurt Pany
Arizona State University

OUR view of the paper by Buckless and Peace (1993) raises questions concerning the actual manipulations of the source of professional standards, the judge's instruction variable, and inferences made by the authors.

I. Source of Professional Standards

The first research hypothesis considers whether the source of professional auditing standards is important in determining juror decisions. Given recent court decisions with increasingly large damage awards against CPA firms, the study's finding that governmental standards result in less frequent decisions against CPA firms would seem to have significant policy implications for CPAs. This finding makes it especially important to understand in detail the nature of the levels of the variables tested in this study.

A careful reading of the two forms of the professional standards used in the study reveals that the variable includes a combination of two manipulations, the type of expert witness and the source of standards. The situation may be viewed as:

	CPA Expert Witness	*Government Official Expert Witness*
Professional Auditing Standards	(1)	(2)
Government Auditing Standards	(3)	(4)

Submitted August 1992.
Accepted September 1992.

Some respondents received the cell (1) combination and others received the cell (4) combination. None, however, received the equally possible cell (2) or (3) treatment, which makes it impossible to separate the source-credibility effects of expert witness from those of the issuer of the standards.

The above analysis suggests that the authors' statement that "The results indicate that jurors attached greater credibility to auditing standards established by the federal government than to standards established by the auditing profession," is inappropriate. It is just as likely that the respondents attached greater credibility to the witness from the government than the witness from the profession.[1]

II. Judge's Instructions

The second manipulation deals with the judge's instructions. Those instructions are:

> *Only Consider Auditing Standards*
>
> "The Judge has instructed you that the professional auditing standards are the standards by which you should assess the defendant's (Adams) performance during the conduct of the audit in question. The Judge has instructed you that Adams should not be held liable for failure to detect fraud unless Adams failed to comply with professional auditing standards."
>
> *Consider Auditing Standards as Part of the Evidence*
>
> "The Judge has instructed you that the professional auditing standards should be taken under consideration by you as a part of the evidence in determining whether the defendant (Adams) exercised the required level of care, skill and diligence during the conduct of the audit in question."

The first form of this manipulation in essence says that the decision should be based only upon professional (or governmental in the alternate case) standards. The second form states that these standards should be taken under consideration as only a part of the evidence. Is it surprising that more would rule against the auditor in the second situation?[2] This issue is especially relevant since the respondents were only told that standards had been met and were not given enough information to make a reasonable judgment concerning the meaning of the meeting of either kind of standard,[3] nor were they given any information on other types of evidence.[4] Because the respondents were given very little information upon which to make their decision concerning the judge's instructions, they were forced to concentrate on the limited information they had. In such a situation, they may react significantly differently to items which otherwise would have been considered minor. The authors' statement that "The results also emphasize the importance of a judge's instruction on the evidentiary value of compliance with auditing standards," is only appropriate within the very narrow context of this study.

---

[1] In addition to the confounding effects of the expert witness variable, it is important to note that the "professional auditing standards" level of the source manipulation does not necessarily summarize the manner in which professional standards are currently issued in the United States and thus limits the generalizations that may be made.

[2] The discussion pertaining to this variable would lead one to expect a test of the interaction between the source and judge instruction variables, as contrasted to the judge instruction variable itself.

[3] It should be noted that an informed reading of the "government auditing standards" level of the source manipulation would probably suggest that the company has had an audit in accordance with "yellow book" standards and the reader would realize that at a minimum, auditor reporting responsibilities under such an audit exceed those of a GAAS-type audit.

[4] The questionnaire included no manipulation or knowledge checks addressing respondent knowledge of the current situation in this country for establishing or meeting either governmental or professional auditing standards. Nor does it describe any additional evidence that may have been useful in their decision.

III. Concluding Remarks

Overall, this study represents an early attempt at considering the responses of jurors to two possible methods of setting auditing standards and of instructing a jury. Because of the potential policy implications of the results of this research, it is important to clearly understand its limitations. For example, the authors' statement that "Jurors' attitudes are affected when they perceive that a standard of performance is self-serving," is inappropriate. There is no evidence that the respondents either considered the standards self-serving or based their verdict on such a belief. In general, it is inappropriate to interpret the results to mean anything beyond that it appears that the source of standards (including both the expert witness and the manner in which standards are established) and the judge's instructions to consider other evidence (even when none is provided) is related to a respondent's finding for the CPA or Plaintiff within the context in which the issue was studied.

Beyond the above comments, the paper highlights several additional limitations. Future research might effectively deal with limitations noted in the paper and suggestions included in this comment.

References

Buckless, F. A. and R. L. Peace. 1993. The influence of the source of professional standards on juror decision making. *The Accounting Review* 68 (January): 164–175.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 179–183

A Comment on "Valuation of Executive Stock Options and the FASB Proposal"

L. Peter Jennergren
Bertil Näslund
Stockholm School of Economics

IN a recent FASB proposal, Swieringa (1987) tentatively concluded that executive stock options (ESOs) should be valued at fair values, with a lower bound constraint. In a recent article in this journal, Foster et al. (1991) argue that a candidate model to establish the fair value of an ESO is the Black-Scholes (B-S) call option pricing model with a constant, continuous dividend yield. The B-S model, however, misses one aspect of ESOs that is different from ordinary call options: ESOs have stochastic lives. That is, an ESO is personal and remains valid only for as long as the executive in question remains in the employ of the company that originally granted the ESO.

The possibility that an ESO will not be exercised leads to a reduction in value, in comparison with an ordinary call option. Foster et al. (1991, 597) point out that this is a "potentially severe" departure from the assumptions of received option theory, and other authors have also noted this feature of ESOs (Noreen and Wolfson 1981, 385–86; Weygandt 1977, 48–49).

The purpose of this comment is to show that ESOs can actually be valued in a fashion that explicitly incorporates their stochastic life feature if one can quantify the rate at which executives leave their current jobs. Our proposed valuation model reduces to the B-S model multiplied by the probability that the ESO will not be cancelled prior to expiration for the particular case of the ESO that can be exercised only upon termination. More complex cases must be handled through numerical methods.

Section I of this comment outlines our valuation model. In section II, we consider the special case of an ESO that may not be exercised before expiration, and in section III, a more complex case that permits exercise before expiration. Section IV contains concluding remarks. Formally, our valution model is a partial differential equation subject to various initial and side conditions (just like the ordinary B-S model). A separate appendix contains the derivation of that partial differential equation.

This comment is concerned with ESO valuation from the *company's* point of view. However, section IV includes a brief remark on valuation from the point of view of *an individual executive.*

Submitted January 1992.
Accepted June 1992.

I. A Valuation Model

We denote by S the current value of the stock of some company, and by $C(S,\tau)$ the value of an ESO on that stock; τ is time to "regular" expiration. The risk-free interest rate is denoted by r, the dividend yield on the stock by q, the stock volatility by σ, and the exercise price by X.

The ESO may be cancelled prematurely, at the rate λ per unit time. The cancellation events, which may also be referred to simply as jumps, follow a Poisson process. It is assumed throughout that λ is constant (i.e., independent of the current stock price S and time to expiration τ). We return briefly to that assumption in the concluding section.

Assume, furthermore, that the jump risk is nonpriced, that is, it can be diversified away. This is not an unreasonable assumption for a large company because the departure of one executive is merely one element of the total set of circumstances that constitute nonsystematic risk (cf. also Mayers and Smith 1986). The value of the ESO, $C(S,\tau)$, then follows the partial differential equation,

$$0.5\sigma^2S^2C_{SS}+(r-q)SC_S-C_\tau-rC+\lambda[K(S,\tau)-C(S,\tau)]=0, \quad (1)$$

where $K(S,\tau)$ is the payoff to the holder of the ESO if there is a jump event (i.e., the ESO is cancelled because the executive leaves the company). This payoff could typically be specified as follows: If the ESO cannot be exercised at the time of the jump event, $K(S,\tau)=0$ (i.e., the ESO becomes worthless). If it can be exercised at that point in time, the executive will presumably arrange to do so before quitting, so $K(S,\tau) =\max(S-X,0)$. The arguments required to derive equation (1) are presented in the appendix.

The initial condition, when specifying the value of the ESO upon "regular" termination (at $\tau=0$), is the usual one,

$$C(S,0)=\max(S-X,0). \quad (2)$$

Also, if exercise is permitted prior to termination, then there must be a side condition stating that exercise will take place, if such action maximizes the value of the ESO to the holder.

II. A Special Case

Suppose that the ESO cannot be exercised before expiration, so that $K(S,\tau)=0$ for all τ. The solution to equations (1) and (2) is then:

$$C(S,\tau)=\exp(-\lambda\tau)\bullet W(S,\tau;X,r,q,\sigma), \quad (3)$$

where $W(S,\tau;X,r,q,\sigma)$ is the B-S value with a constant, continuous dividend yield of q (see appendix).

The first factor in equation (3) is recognized as the probability of zero jumps (from the Poisson distribution) during the lifetime, τ, of the ESO. Hence, we have a very intuitive result: The value of an ESO, which may become worthless if the executive leaves the company, equals the ordinary B-S value multiplied by the probability that the executive will stay with the firm until the ESO expires. The difference between the B-S value and that given by equation (3) may not be negligible. For instance, if the current share price $S=100$, the time to expiration $\tau=3$ years, the exercise price $X=100$, the risk-free interest rate $r=0.08$, the dividend yield $q=0.03$, and the stock volatility $\sigma=0.3$, then the B-S value is 24.50. However, if the jump rate λ is 0.05 per year, then the ESO value falls to 21.09.

Formula (3) is valid only if the ESO cannot be exercised before expiration. A difference between an ESO and an ordinary call option is that it may be optimal to exercise an ESO before expiration (if that is permitted), even if the underlying stock pays no dividends (i.e., even if q equals zero); this is not so for an ordinary call option. Also, the value of an ESO on a nondividend stock may not increase with increasing time to expiration, unlike ordinary call options on such stocks.

III. A More Typical ESO Plan

Consider now a more typical case. Suppose that the time to expiration of some ESO is ten years. During the first three years the ESO cannot be exercised. After that, exercise is permitted. Further assumptions are as follows:

$S=100$ (current share price),
$X=100$ (exercise price),
$r=0.08$ (risk-free interest rate),
$q=0.03$ (dividend yield on the stock), and
$\sigma=0.3$ (stock volatility).

As before, λ is the rate per year for the ESO to be cancelled because the executive leaves the company. Since the ESO can be exercised during the final seven years of the contract, but not during the first three, $K(S,\tau)$ is as in section I, that is,

$$\begin{aligned} K(S,\tau)&=0 && \text{for } 10\geq\tau>7, \\ K(S,\tau)&=\max(S-X,0) && \text{for } 7\geq\tau>0. \end{aligned}$$

The initial condition at $\tau=0$ is equation (2). Finally, since the ESO can be exercised after three years, it must hold that,

$$C(S,\tau)\geq S-X \qquad \text{for } 7\geq\tau>0.$$

The value of the ESO can now be calculated by means of a numerical method (our preferred choice is the implicit finite difference method; see Hull 1989, 234–42). If $\lambda=0$, then the resulting value for one ESO is 40.51. For $\lambda=0.02$, the corresponding value is 37.31. For $\lambda=0.05$, the value is 33.07. Even a fairly small annual risk of the executive's leaving can bring about a noticeable drop in ESO value.

IV. Conclusion

The ordinary B-S formula and extensions to incorporate exercise prior to expiration (when that is permitted) overstate the value of an ESO by disregarding the fact that ESOs are really options with stochastic lives. The stochastic life feature of an ESO reduces its value compared to that of an ordinary call option, since there is a possibility that it will be cancelled prematurely at conditions less favorable to the holder than regular termination at the expiration of the contract.

If ESOs are to be valued by a single formula, then we suggest formula (3) above, rather than the B-S formula. More realistic cases, in particular those that permit exercise prior to expiration, can be handled through fairly simple numerical procedures, as illustrated in section III.

When valuing an ESO, there is evidently one more parameter to estimate, λ, the rate at which an executive may leave his or her job. If λ is constant (i.e., does not depend on the current stock price S and time τ to expiration), as we have assumed, then it should not be impossible to obtain rough estimates of that parameter (on the basis of recent experiences in companies granting ESOs). However, λ could depend on S and τ, as also suggested by Noreen and Wolfson (1981, 386). We have not considered this possibility because we do not have any clear idea about a plausible relationship between λ, on the one hand, and S and τ, on the other hand. For instance, are executives more likely to switch jobs when the share prices of their companies are low or high? This seems to be an interesting area for further research.

Finally, it is worth emphasizing that our suggested valuation model is applicable to the value of an ESO only from the point of view of the issuing company. To the issuer, an ESO is an option with a stochastic life, which reduces its value compared to an ordinary call option. From the point of view of the holder, an ESO has the additional quality of being highly illiquid, since it cannot be traded. This lack of liquidity means a further reduction in value to the holder (but not to the issuing company; cf. also the concluding paragraph of the appendix). Moreover, the assumption that the jump risk can be diversified away, which is not unreasonable when applied to the issuing company, is not realistic when applied to the holder. The value of an ESO for the holder can be expected to be lower than its value for the issuer, as also suggested by Lambert et al. (1991, 135–36).

Appendix

The derivation of the differential equation (1) is similar to the procedure in Merton (1976), with the important difference that here it is the *option, not the underlying stock*, that is subject to jump events. As stated above, jumps are assumed to follow a Poisson process with jump rate λ per unit time. Let the stock return dynamics be given by:

$$dS/S=(\alpha-q)dt+\sigma dz,$$

where dz is a standard Wiener process. Here, $C(S,t)$ denotes the value of the ESO as a function of chronological time, t, rather than time to expiration, τ [and similarly for $K(S,t)$].

By Ito's lemma, and by the assumption that the option may jump to a value of $K(S,t)$, the return dynamics of the option value $C(S,t)$, are given by:

$$dC/C=\{[0.5\sigma^2S^2C_{ss}+(\alpha-q)SC_s+C_t]/C\}dt+(\sigma SC_s/C)dz$$
if the jump does not occur,

$$dC/C=\{[0.5\sigma^2S^2C_{ss}+(\alpha-q)SC_s+C_t]/C\}dt+(\sigma SC_s/C)dz+[K(S,t)-C(S,t)]/C(S,t)$$
if the jump occurs,

where C_s, C_{ss}, and C_t are the usual notations for partial derivatives of C with respect to S and t. We define,

$$\alpha_C=\{[0.5\sigma^2S^2C_{ss}+(\alpha-q)SC_s+C_t+\lambda(K-C)]/C\}, \tag{1A}$$

$$\sigma_C=[\sigma SC_s/C], \tag{2A}$$

and note that α_C is the expected rate of return on the option.

We now compose a portfolio of the stock, the option, and the risk-free asset. Let the portfolio proportions be x_1 (stock), x_2 (option), and x_3 (risk-free asset), with $x_1+x_2+x_3=1$. Because the proportions sum to unity, the resulting portfolio expected rate of return is,

$$x_1[(\alpha-q)+q-r]+x_2(\alpha_C-r)+r.$$

Note that x_1 and x_2 are chosen to eliminate the risk associated with the movement of the *stock* price, so that.

$$x_1\sigma+x_2\sigma_C=0.$$

The remaining risk associated with the portfolio is hence only the option jump risk. But suppose that the jump risk is nonpriced (i.e., can be diversified away). If so, the expected return on the portfolio must be the risk-free rate of return r. This leads to the equation,

$$[(\alpha-q)+q-r]/\sigma=(\alpha_C-r)/\sigma_C. \tag{3A}$$

Substituting equations (1A) and (2A) into equation (3A), and changing from chronological time, t, to time to expiration, τ, we obtain the partial differential equation (1) in section I for $C(S,\tau)$.

If $K(S,\tau)=0$, then equation (1) can be written as,

$$0.5\sigma^2S^2C_{SS}+(r-q)SC_S-C_\tau-(r+\lambda)C=0. \tag{4A}$$

Defining $r^0=(r+\lambda)$, we can rewrite equation (4A) as,

$$0.5\sigma^2S^2C_{SS}+[r^0-(q+\lambda)]SC_S-C_\tau-r^0C=0.$$

Using the formula for the B-S call option value with a constant, continuous dividend yield (see, e.g., Hull 1989, 135), and using the fact that $r^0=(r+\lambda)$, one obtains equation (3).

The personal nature of an ESO, which means that it cannot be traded, implies that the holder cannot carry out the continuous arbitrage required in this derivation. However, this does not invalidate the valuation equation (1). If equation (1) does not hold, profitable arbitrage would be possible by some market participant other than the holder of the ESO, for instance by the company that granted the ESO. But no profitable arbitrage can be possible in equilibrium (Ingersoll 1987, 323; cf. also Noreen and Wolfson 1981, 386).

References

Foster, T. W. III, P. R. Koogler, and D. Vickrey. 1991. Valuation of executive stock options and the FASB proposal. *The Accounting Review* 66 (July): 595–610.

Hull, J. 1989. *Options, Futures, and Other Derivative Securities.* Englewood Cliffs, NJ: Prentice-Hall.

Ingersoll, J. E., Jr. 1987. *Theory of Financial Decision Making.* Totowa, NJ: Rowman & Littlefield.

Lambert, R. A., D. F. Larcker, and R. E. Verrecchia. 1991. Portfolio considerations in valuing executive compensation. *Journal of Accounting Research* 19 (Supplement): 129–49.

Mayers, D., and C. Smith. 1986. The corporate insurance decision. *The Revolution in Corporate Finance.* Oxford, UK: Basil Blackwell, 230–238.

Merton, R. C. 1976. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3: 125–44.

Noreen, E., and M. Wolfson, 1981. Equilibrium warrant pricing models and accounting for executive stock options. *Journal of Accounting Research* 29 (Autumn): 384–98.

Swieringa, R. J. 1987. *Accounting for Stock Options: FASB Status Report No. 183.* Stamford, CT: FASB.

Weygandt, J. J. 1977. Valuation of stock option contracts. *The Accounting Review* 52 (January): 40–51.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 184–189

Valuation of Executive Stock Options and the FASB Proposal: An Extension

Taylor W. Foster III
New Mexico State University
Paul R. Koogler
Southwest Texas State University
Don Vickrey
Arizona State University West

IN Foster et al. (1991), we use the Black and Scholes option-pricing model to estimate the value of executive stock options (ESOs) in a study on the materiality of ESO compensation expense. In this issue of *The Accounting Review*, Jennergren and Näslund (1993) address a potentially important issue regarding the materiality of ESO-related expense. If the employment of the option recipient is terminated, ESOs may be either forfeited or exercised prematurely. The Black and Scholes continuous-dividend model used in Foster et al. relies on four parameters: the exercise price, the variance rate of return on the optioned stock, a risk-free interest rate, and a continuous-dividend rate. Jennergren and Näslund's valuation procedure allows for the possibility of premature cancellation by adding one additional parameter, λ, which is defined as the rate per unit time that ESOs are cancelled prematurely. They modify the Black and Scholes model for λ, thus rendering it appropriate for ESOs that are not exercisable until maturity (i.e., European options). The central concern, however, is the "more typical plan" for ESOs, those that are exercisable prior to maturity (i.e. American options). Jennergren and Näslund suggest a numerical method, adjusted for λ, to estimate the value of such options.

In concept, λ must be constant through time, follow a Poisson process (i.e., apply to a large group of employees), and be fully diversifiable. In practical applications, λ must also be observable. We believe that λ is difficult to estimate and is unstable through time. Moreover, we suspect that λ is unique to each employee position and possibly to each employee. Such issues must be resolved empirically before λ could affect accounting policy. Finally, we feel that λ and the related numerical method are likely to be incompatible with the Financial Accounting Standard Board's (FASB) policy-setting precedents.

Invited by the Editor.
Accepted September 1992.

Notwithstanding λ's likely instability, and so forth, we implement Jennergren and Näslund's (1993) suggestion and reassess the materiality of ESO compensation expense, assuming grant-date accounting and a five-year amortization period. We apply the FASB's proposal, but use assumed values of λ and a numerical procedure in place of the Black and Scholes model to estimate the fair value of an ESO. We assume values for λ of 0.00, 0.02, and 0.05. In addition, we use λ and the numerical procedure without the FASB's minimum-value constraint to assess the materiality of ESO compensation expense. The FASB proposal, augmented by a λ less than or equal to 0.05, yields essentially the same percentage of firms with material ESO compensation as reported in Foster et al. (1991). Estimation using λ without the minimum-value constraint still generates a substantial percentage of firms in our sample with material ESO compensation expense.

The remainder of our response is organized as follows. Section I gives the reasons for our belief that λ is unstable and difficult to estimate reliably. Section II discusses the compatibility of Jennergren and Näslund's (1993) suggested procedure with the formulation of accounting policy. Section III compares the percentage of firms that would report material ESO compensation expense, using Jennergren and Näslund's procedure, with selected percentages reported in Foster et al. (1991). Concluding remarks comprise section IV

I. Stability of the Added Parameter (λ)

The magnitude of λ should depend on numerous conditions, including the mobility and promotion potential of the employee as well as his or her propensity to change employment. These conditions, in turn, depend, for example, on the employee's skills and the transferability of these skills; the employee's utility function, wealth, and probability beliefs; the employee's position within the firm; the employee's age; the industry in which the firm operates; employment alternatives; and the general economy. Many of these conditions are likely to change through time. We suspect that the intertemporal dynamics of these conditions make λ unstable and difficult to estimate reliably.

Since other parameters (risk-free rate, variance rate, and dividend rate) might also be unstable and are nevertheless assumed constant in option valuation, instability by itself does not justify ignoring λ in estimating compensation related to ESOs. Thus, we now direct attention to issues and results related to implementing λ and the numerical procedure.

II. The Compatibility of λ and the Numerical Method with Accounting Policy

The suggestion by Jennergren and Näslund (1993) is good, in concept, despite the likely instability of λ. However, we doubt that the FASB would adopt their procedure for three reasons. First, the FASB's prescibed accounting for market-based compensatory plans is to treat employee turnover as a change in accounting estimate (i.e., prospectively). For example, Interpretation No. 28, which applies to stock appreciation rights, requires that "accrued compensation for a right that is forfeited or cancelled shall be adjusted by decreasing compensation expense in the period of forfeiture . . ." (FASB 1981a, par. 4). Similarly, the extant treatment of compensation expense related to ESOs does not allow for anticipated turnover. Thus, the FASB has opted to ignore turnover in relation to market-based compensation plans until it becomes a reality.

Second, the Board would have to provide specific guidance on the implementation of a numerical method with λ. Estimates of ESO values are almost as sensitive to how the numerical procedure is carried out (choice of maximum stock price, stock price increments, and time intervals) as they are to variations in λ. Our experience with the numerical procedure indicates that ESO estimates are quite sensitive to choices of maximum stock price and stock price increments. Thus, we believe the FASB is unlikely to become involved in these implementation issues.

Third, the FASB has been reluctant to impose excessive accounting costs on firms, especially if the more costly and accurate method is unlikely to make a material difference in accounting results.[1] We predict two types of firm-related implementation costs. First, firms would incur costs to attain reasonably accurate and auditable estimates of λ. Management would have to track turnover, possibly by employee classification, within the company and possibly within the industry.[2] Reliable turnover data, especially for high-ranking executives and directors, might be unavailable. Provided turnover data are available or tracking is possible, turnover estimates would then have to be translated into rates per unit time. Additionally, costly actuarial methods might have to be employed to estimate λ. Second, as we indicated above, ESO estimates are sensitive to the particular numerical procedure used (e.g., implicit vs. explicit finite differencing) and to choice of maximum stock price, stock price increments, and time intervals. Thus, existing personnel would have to learn numerical methods, write related programs, and implement the numerical procedure. Alternatively, the firm would have to hire new employees or pay consultants to perform these tasks.[3] We doubt that the FASB will require firms to incur such costs in estimating λ and implementing numerical procedures, particularly if these refinements make little difference in the amount of ESO compensation expense reported. As shown below, the use of λ and the numerical procedure is very unlikely to appreciably reduce ESO compensation expense.

III. ESO Estimation Revisited

In this section, we report the percentage of sample firms in Foster et al. (1991) that would report material ESO compensation expense using λ with the numerical procedure suggested in Jennergren and Näslund (1993) in place of the Black and Scholes model.[4] First, we use λ and the numerical procedure in the FASB proposal. We assume the same values of λ as Jennergren and Näslund, 0.00, 0.02, and 0.05.[5] Recall that the

---

[1] SFAS No. 52, e.g., allows the use of the weighted-average exchange rate instead of the prevailing spot rate in translating income statement items (FASB 1981b).

[2] A pertinent classification scheme might include technical staff, junior executives, senior executives, inside directors, and outside directors. However, if λ is indeed unique to each employee, a possibility noted earlier, it cannot be estimated and our discussion of estimation costs and related issues is irrelevant.

[3] Numerical methods rely on a finite differencing technique to analyze a partial differential equation by using discrete estimates of changes in the option's value for changes in time or the optioned stock price. Finite difference techniques thus form difference equations as approximations to continuous partial derivatives. There are an infinite variety of ways to estimate the changes in an option's value with respect to time and stock price. Forward, central, and backward differences can be used, as well as complex averages of them (Geske and Shastri 1985, 51–52). Thus, there are a large number of numerical procedures that could be implemented, but it is unclear which procedure is optimal for ESOs.

[4] We use the numerical procedure suggested by Jennergren and Näslund, the implicit finite difference method (see Hull 1989, 234–42).

[5] According to Business and Legal Reports, Inc. (BLR 1991), several factors affect turnover rates and imply that separation rates are unreliable, especially for growing firms and declining firms whose separation rates

FASB proposal requires that ESOs be valued at fair value, but at not less than a minimum-value constraint (Foster et al. 1991, 597). We also consider the materiality of ESO compensation expense using ESO estimates from the numerical procedure without the minimum-value constraint.

In both cases, we assume grant-date accounting (see Foster et al. 1991, 606–07) and a five-year amortization period. A five-year service period is used because it is a representative period that is common in the sample of firms in Foster et al. To avoid specific grant-date assumptions and fractional-year amortization, we estimate ESO compensation expense for fiscal year 1986 (see fn. 23 in Foster et al. 1991, 604). Fiscal year 1986 is a representative year from the five-year service period. We divide compensation expense for each ESO grant by the absolute value of operating income for fiscal year 1986 and compare this result with materiality criteria of 3 and 5 percent.

In table 1, we report the respective percentages of material ESO compensation expense, obtained by augmenting the FASB proposal with λ, for dividend-paying ($n=170$) and nondividend firms ($n=44$). The percentage for the 5 percent criterion is in parentheses. If λ is 0.00, 29 (21) percent of dividend-paying firms would disclose a material charge to operating income. In Foster et al. (1991), we report 29 (22) percent. A λ of 0.02 somewhat diminishes the percentage, to 29 (20); a λ of 0.05 further diminishes the percentage to 28 (20). A λ equal to 0.00 yields material ESO compensation expense for 52 (50) percent of the nondividend firms. In Foster et al., we report 50 (48) percent. If λ is 0.02 or 0.05, the percentage is 50 (48). Thus, the use of λ, in combination with the FASB proposal, only negligibly affects the materiality of ESO compensation expense for the sample firms in Foster et al., even when using relatively high λs (see fn. 5).

The negligible effect of λ in the FASB proposal, relative to the results in Foster et al. (1991), is partially a consequence of the FASB's minimum-value constraint. The percentage of ESO grants for which minimum value exceeds fair value, also reported in table 1, gives additional perspective on the effect of λ. For dividend-paying firms, λ-related estimates exceed the corresponding minimum-value estimates in 47, 31, and 21 percent of the cases for respective values of λ equal to 0.00, 0.02, and 0.05. In Foster et al. (1991), we report that Black and Scholes continuous-dividend estimates exceed minimum-value estimates for 46 percent of the firms. For nondividend firms, the λ-related estimates exceed the minimum-value estimates in 100, 64, and 48 percent of the cases for respective values of λ equal to 0.00, 0.02, and 0.05.[6]

"fluctuate wildly." In addition, no particular percentage of turnover is normal for all companies. Turnover varies with company size, type of employee, and time of year. In general, turnover rates decrease as the number of employees in a given firm increases. According to BLR, companies with less than 250 employees experienced an average turnover rate of 1.1 percent for 1989; companies with more than 2,500 employees experienced an average turnover of 0.8 percent. The 1989 average for all companies is 1.1 percent. BLR reports that the average turnover for 1989 by industry is as follows: manufacturing, 0.8 percent; nonmanufacturing, 1.3 percent; finance, 1.5 percent; nonbusiness, 1.2 percent; healthcare, 1.5 percent. We note that ESO recipients represent a subset of the employees covered by these turnover rates, and the ESOs purportedly tie the recipient to the firm. Thus, the turnover rates for ESO recipients presumably are much lower than those reported above by industry and lower than the positive rates used here and in Jennergren and Näslund (i.e., 2 percent and 5 percent).

[6] The Black and Scholes continuous-dividend model reduces to the ordinary Black and Scholes model if the firm does not distribute cash dividends. For nondividend firms, numerical method estimates with $\lambda=0.00$ should be approximately equal to ordinary Black and Scholes estimates; both estimates should be at least as great as minimum-value estimates. In Foster et al. (1991) the ordinary model gives ESO estimates that are greater than those from the minimum-value model for all nondividend firms.

Table 1
The Effect of Premature Cancellation Rates, λs, on the Materiality of ESO Compensation Expense

	$\lambda=0.00$	$\lambda=0.02$	$\lambda=0.05$	*Results from Foster et al.* (1991)
Material expense[a] with FASB proposal[b]				
Dividend-paying firms	29% (21%)	29% (20%)	28% (20%)	29% (22%)
Nondividend firms	52% (50%)	50% (48%)	50% (48%)	50% (48%)
Fair value > minimum value[c]				
Dividend-paying firms	47%	31%	21%	46%
Nondividend firms	100%	64%	48%	100%
Material expense,[a] with numerical method				
Dividend-paying firms	26% (20%)	22% (18%)	21% (18%)	nr[d]
Nondividend firms	48% (48%)	48% (48%)	48% (43%)	nr[d]

[a] The percentages represent the firms with material ESO compensation expense when using the 3% (5%) criterion. There are 170 dividend-paying firms and 44 nondividend firms.

[b] Under the FASB proposal, the value of ESOs = max(fair value, minimum value).

[c] The percentage of firms for which the ESO value from the numerical method > the ESO value from the minimum-value model. The results reported for Foster et al. (1991) show the percentage of firms for which the ESO value from the Black and Scholes model > the ESO value from the minimum-value model.

[d] Not reported in Foster et al. (1991).

Table 1 also shows the percentage of firms with material ESO compensation for fiscal year 1986 when using the numerical method without the minimum-value constraint. If λ equals 0.00, 26 (20) percent of the dividend-paying firms would disclose material ESO compensation expense. A λ equal to 0.02 reduces the percentage to 22 (18), and λ equal to 0.05 further reduces the percentage to 21 (18). If λ equals 0.00 or 0.02, 48 (48) percent of the nondividend firms would report material ESO compensation. A λ of 0.05 somewhat reduces the percentage to 48 (43). Thus, in the absence of the minimum-value constraint, a relatively large percentage of the sample firms in Foster et al. (1991) would disclose material charges to operating income when λ is utilized.[7] These percentages are similar to the results implicit in Foster et al. (1991), especially for λ equal to 0.02. They are not drastically different even for λ equal to 0.05.

---

[7] We programmed the numerical method for a three-year vesting period. Assuming an ESO with a life of ten years, a three-year vesting period dictates a boundary condition of zero for the ESO's payoff during the first three years after the grant date; thereafter, for the remaining seven years until expiration, the boundary condition for the ESO's payoff is $\max(0, S-X)$, where S is the current price of the optioned stock and X is the exercise price. Had we assumed a two-year vesting period, the boundary condition for the ESO's payoff would have been zero for the first two years after grant date and $\max(0, S-X)$ thereafter until the ESO expires. For any $\lambda > 0$, the numerical procedures should give a greater value for an ESO with a two-year vesting period than for an identical ESO with a three-year vesting period. Hence, ESO-related compensation estimates, for $\lambda > 0$, used to calculate the materiality percentages reported herein are less than what would obtain had we assumed a two-year vesting period. According to Foster et al. (1991, 607), a survey of CFOs indicates an average vesting period

IV. Summary

The suggestion by Jennergren and Näslund (1993) is interesting in the context of estimating compensation expense related to ESOs. However, we find practical problems in implementing the numerical procedure to estimate ESO compensation expense. ESO estimates from the numerical procedure are sensitive to how the program is set up (e.g., choice of maximum stock price, stock price increments, and time intervals). We also suspect problems with the stability and reliability of λ estimates. Several conditions (i.e., mobility, promotion potential, and propensity to change employment) affect the magnitude of λ, and these conditions are likely to change over time. Thus, we believe λ is inherently unstable. If λ is indeed unstable, then estimates of λ based upon historical data are likely to be inaccurate (i.e., recent turnover experiences in companies granting ESOs would not give relevant estimates of λ).

Our results suggest that using λ in the FASB proposal fails to reduce the materiality of ESO compensation expense appreciably. The only potential for a perceptible reduction in ESO compensation estimates appears to occur with relatively high, and possibly unrealistic, values of λ and without the minimum-value constraint.[8] As indicated earlier, the use of λ may be incompatible with the FASB's policy-setting precedents for a variety of reasons, especially given that the effect of λ on the materiality of ESO compensation expense appears negligible.

---

of 1.9 years for ESOs. Thus, the percentages we report in this section are less than would be reported for a more representative vesting period (e.g., two-years). In other words, we believe that λ-related ESO compensation estimates for a representative vesting period will give essentially the same results (i.e., percentage of firms with material ESO compensation expense) as reported in Foster et al. (1991), even if the numerical method is applied without the minimum-value constraint. A time constraint precluded a sensitivity analysis for variations in assumed vesting periods.

[8] The vast majority of ESOs are granted to key employees, individuals who are influential in setting objectives, developing strategy and long-term planning for their employer. Stock options are generally recognized as an incentive to keep such employees with the firm and to extend their planning horizons. In other words, one purpose of stock options is to reduce the turnover of key employees (i.e., reduce λ). We believe that it is very unlikely that a key employee would choose to resign during the time period that ESOs are outstanding. Thus, we also believe that, for key employees, a λ of even 0.02 is likely to be unrealistically high.

References

Business & Legal Reports, Inc. 1991. *Turnover*. Madison, CT: BLR.

Financial Accounting Standards Board. 1981a. *Interpretation No. 28: Accounting for Stock Appreciation Rights and Other Variable Stock Option or Award Plans*. Stamford, CT: FASB.

——. 1981b. *Statement of Financial Accounting Standards No. 52: Foreign Currency Translation*. Stamford, CT: FASB.

Foster, T. W., III, P. R. Koogler, and D. Vickrey. 1991. Valuation of executive stock options and the FASB proposal. *The Accounting Review* 66 (July): 595–610.

Geske, R., and K. Shastri. 1985. Valuation by approximation: A comparison of alternative option valuation techniques. *Journal of Financial and Quantitative Analysis* 20 (March): 45–71.

Hull, J. 1989. *Options, Futures, and Other Derivative Securities*. Englewood Cliffs, NJ: Prentice-Hall.

Jennergren, L. P., and B. Naslund. 1993. A Comment on "Valuation of executive stock options and the FASB proposal." *The Accounting Review* 68 (January): 179–183.

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 190–200

American Accounting Association

Committees 1992–93

EXECUTIVE COMMITTEE

President—Gary L. Sundem, University of Washington
President-Elect—Andrew D. Bailey, Jr., University of Arizona
Vice President—Andre Zund, University of St. Gallen
Vice President—Gary K. Meek, Oklahoma State University
Vice President—Bernard J. Milano, KPMG Peat Marwick
Secretary-Treasurer—Jane O. Burns, Texas Tech University
Director of Education—Robert Ingram, University of Alabama
Director of Research—Katherine Schipper, University of Chicago
Director of Publications—Kenneth J. Euske, Naval Postgraduate School
Past President—Arthur R. Wyatt, University of Illinois at Urbana-Champaign

COUNCIL

Section and Group Representatives

Accounting Behavior & Organizations Section—Mark E. Haskins, University of Virginia, and Paul H. Munter, University of Miami
Administrators of Accounting Programs Group—Michael A. Diamond, University of Southern California
American Taxation Association Section—John O. Everett, Virgina Commonwealth University, and Anna C. Fowler, University of Texas
Auditing Section—William F. Messier, Jr., University of Florida, and Theodore J. Mock, University of Southern California
Gender Issues in Accounting—Rita Hull, Virginia Commonwealth University
Government & Nonprofit Section—Earl R. Wilson, University of Missouri-Columbia
International Section—Frederick D. Choi, New York University, and Kathleen Sinning, Western Michigan University
IS/MAS Section—Anita Sawyer Hollander, Florida State University
Management Accounting Section—Donald L. Madden, University of Kentucky, and L. Gayle Rayburn, Memphis State University
Public Interest Section—Marilyn Neimark, CUNY-Baruch College
Two-Year College Section—Patrick T. Reihing, Nassau Community College

Regional Group Representatives

Mid-Atlantic Regional Group—Larry G. Singleton, George Washington University
Midwest Regional Group—Thomas Hubbard, University of Nebraska, and Gerald Smith, University of Northern Iowa
Northeast Regional Group—Mark Higgins, University of Rhode Island, and Philip Jagolinzer, University of Southern Maine
Ohio Region—Gyan Chandra, Miami University
Southeast Regional Group—Ronald L. Clark, University of Georgia, and Thomas C. Turner, University of North Carolina
Southwest Regional Group—Kent Royalty, St. Mary's University
Western Regional Group—Robert Capettini, San Diego State University, and James F. Sepe, Santa Clara University

Members-at-Large

Julie Collins, University of North Carolina
Jack L. Krogstad, Creighton University
W. Morley Lemon, University of Waterloo
Douglas Snowball, University of Florida

REGIONAL VICE PRESIDENTS

Mid-Atlantic

Larry G. Singleton, George Washington University

Midwest

Gerald Smith, University of Northern Iowa

Northeast

Philip Jagolinzer, University of Southern Maine

Ohio

Gyan Chandra, Miami University

Southeast

Ronald L. Clark, University of Georgia

Southwest

Kent Royalty, St. Mary's University

Western

Robert Capettini, San Diego State University

SECTION OFFICERS

Accounting, Behavior & Organizations Section

Chairperson: Paul H. Munter, University of Miami
Chair-Elect: Don W. Finn, Texas Tech University
Secretary-Treasurer: Lawrence Ponemon, Babson College
Immediate Past Chair: Mark E. Haskins, University of Virginia

Administrators of Accounting Programs Group

President: Michael A. Diamond, University of Southern California
Vice President: Robert G. May, University of Texas at Austin
Secretary: Jerry E. Trapnell, Clemson University
Treasurer: Dora R. Herring, Mississippi State University

American Taxation Association Section

President: John O. Everett, Virginia Commonwealth University
Vice President: Debra W. Hopkins, University of Texas-Arlington
President-Elect: Anna C. Fowler, University of Texas
Secretary-Treasurer: Richard P. Weber, Michigan State University

Artificial Intelligence/Expert Systems

Chairperson: Dan O'Leary, University of Southern California
Vice Chairperson: Rien Bouwman, University of Arkansas
Chairperson-Elect: Carol Brown, Oregon State University
Secretary-Treasurer: Paul Watkins, University of Southern California

Auditing Section

Chairperson: William F. Messier, Jr., University of Florida
Vice Chairperson (Academic): W. Robert Knechel, University of Florida
Vice Chairperson (Practice): Steve Aldersley, Ernst & Young
Secretary-Treasurer: Jane F. Mutchler, Penn State University

Gender Issues in Accounting Section

Chairperson: Rita Hull, Virginia Commonwealth University
Vice Chairperson: Camille Koehler, University of New Mexico
Chairperson-Elect: Theresa Hammond, Boston College
Secretary-Treasurer: Karen McCarron, Georgia State University

Government and Nonprofit Section

Chairperson: Earl R. Wilson, University of Missouri-Columbia
Chairperson-Elect: W. T. Wrege, Ball State University
Secretary: Robert W. Parry, Jr., Indiana University

Information Systems/Management Advisory Services Section

Chairperson: Anita Sawyer Hollander, Florida State University
Vice Chairperson (Academic): Paul J. Steinbart, Memphis State University
Vice Chairperson (Practice): Rod C. Thomas, Coopers & Lybrand
Secretary-Treasurer: Ronald L. Clark, University of Georgia

International Accounting Section

Chairperson: Kathleen E. Sinning, Western Michigan University
Vice Chairperson (Academic): Frederick D. Choi, New York University
Vice Chairperson (Practice): Keith Scott, Certified General Accountants Association of Canada
Secretary: James A. Schweikert, University of Richmond

Management Accounting Section

President: L. Gayle Rayburn, Memphis State University
President-Elect: Michael A. Robinson, Baylor University
Secretary-Treasurer: Susan F. Haka, Michigan State University
Past President: Donald L. Madden, University of Kentucky

Public Interest Accounting Section

Chairperson: Marilyn Neimark, CUNY-Baruch College
Vice Chairperson: Robert G. Ruland, Suffolk University
Secretary-Treasurer: Paula B. Thomas, Middle Tennessee State University
Chair-Elect: Mary S. Doucet, University of Georgia

Teaching and Curriculum Section

Chairperson: Jan R. Williams, University of Tennessee
Vice Chairperson (Academic): Jay M. Smith, Jr., Brigham Young University
Vice Chairperson (Practice): Dennis R. Reigle, Arthur Andersen & Co.
Secretary: E. Kent St. Pierre, James Madison University
Treasurer: Mary Beth Armstrong, California Polytechnic State University

Two Year College Section

Chairperson: Patrick T. Reihing, Nassau Community College
Vice Chairperson: Linda Lessing, SUNY-Farmingdale
Secretary-Editor: Ellen Sweatt, Dekalb College
Coordinator of Regional Representatives: Billie Cunningham, Collin County Community College

ACCOUNTING ACCREDITATION COMMITTEE

Charges: 1. To monitor the AACSB review of accreditation and other developments in accreditation that would potentially impact accounting programs.

2. To inform the AAA Executive Committee of relevant accreditation issues and recommend policies or positions as appropriate.

Chairperson: David R. Campbell, Case Western Reserve University
Deborah F. Beard, Southeast Missouri State University
Lanny G. Chasteen, Oklahoma State University
James A. Heintz, University of Connecticut
Joseph A. McHugh, University of Massachusetts at Boston
Alan Reinstein, Wayne State University
Michael G. Tearney, University of Kentucky

ACCOUNTING EDUCATION ADVISORY COMMITTEE

Charges: 1. To develop and issue position statements on matters relating to accounting education.

2. To develop and recommend to the Executive Committee educational policy for the Association.

3. To develop and present to the Executive Committee a strategic plan for the Association's role in professional development.

4. To initiate, coordinate, and administer the activities of the Association in the field of accounting education, including accounting education research and continuing professional education. Effective liaison with regions and sections is an integral part of this responsibility.

5. To coordinate the work of committees for the Doctoral Consortium, Doctoral Programs, New Faculty Consortium, Senior Faculty Consortium, Trueblood Seminars, Professional Examinations, and other education-designated committees that may be established during the year.

Chairperson: Robert W. Ingram, University of Alabama
Bruce A. Baldwin, Arizona State University-West Campus
Elba F. Baskin, Arthur Andersen & Co.
Billie Cunningham, Collin County Community College
Gary L. Fish, Illinois State University
Karen L. Hooks, University of South Florida
Paul A. Janell, Northeastern University
Russell J. Petersen, University of Akron
David N. Ricchiute, University of Notre Dame
Shyam Sunder, Carnegie Mellon University
Jan R. Williams, University of Tennessee

ACCOUNTING EDUCATION CHANGE LIAISON COMMITTEE

Charges: 1. To monitor developments of the Accounting Education Change Commission.

2. To consider alternative initiatives for the AAA that are responsive to AECC developments.

3. To recommend to the Executive Committee specific initiatives for AAA implementation.

Chairperson: Donald E. Kieso, Northern Illinois University
William F. Bentz, Ohio State University
Patricia A. Cook, Manchester Community College
Brent Inman, Coopers & Lybrand
Gary A. Luoma, University of South Carolina
Rene P. Manes, Florida State University
Patricia Metzger, Mary Washington College

ASSOCIATION STRUCTURE EVALUATION FOLLOW-UP COMMITTEE

Charge: To follow-up on the recommendations of the 1991–1992 AAA Structure Evaluation Committee, considering the feedback on their report at the April 1992 Council and Executive Committee meetings. Areas of focus should include, but are not limited to:

1. A structure to more effectively bring AAA services to a larger number of members.
2. The structure, operations, and membership of Council.
3. The membership and selection of the Executive Committee, including Nominating Committee membership, selection, and procedures. Specific recommendations should be submitted to the Executive Committee by March 1, 1993.

Chairperson: William L. Felix, Jr., University of Arizona
Alvin A. Arens, Michigan State University
Kathleen R. Bindon, University of Alabama
W. Morley Lemon, University of Waterloo
Charles J. Pineno, Clarion University
Patrick Reihing, Nassau Community College
Howard O. Rockness, University of North Carolina
Mark A. Wolfson, Stanford University

BY-LAWS COMMITTEE

Charge: To review the Association's By-Laws and recommend appropriate revisions to the Executive Committee.

Chairperson: Corine T. Norgaard, University of Connecticut
Jane O. Burns, Texas Tech University
Marcos F. Massoud, Claremont McKenna College
Jacqueline B. Sanders, Mercer County Community College
S. Sam Sedki, St. Mary's University

CARTER SCHOLARSHIPS COMMITTEE

Charges: 1. To collect and evaluate applications for the Carter Foundation Student Scholarships.
2. To recommend amounts of stipends.
3. To rank applications and to transmit the ranking to the trustees of the Foundation.

Chairperson: Geraldine F. Dominiak, Texas Christian University
John Cumming, Miami University
Wayne E. Leininger, Virginia Poly Institute and State University

COMMITTEE ON NOMINATIONS

Charge: To select a list of nominees to AAA offices for election by the membership in August 1993.

Chairperson: John K. Simmons, University of Florida
Alvin A. Arens, Michigan State University
Julie H. Collins, University of North Carolina
Fredrick L. Neumann, University of Ilinois
Edward J. Schnee, University of Alabama
Carl S. Warren, University of Georgia
Arthur R. Wyatt, University of Illinois at Urbana-Champaign

COMPETITIVE MANUSCRIPT AWARD COMMITTEE

Charge: To select one, two, or, at most, three recipient(s) of the Competitive Manuscript Award using criteria approved by the Executive Committee.

Chairperson: David Burgstahler, University of Washington
Stanley Baiman, University of Pennsylvania
Srikant M. Datar, Stanford University
John W. Dickhaut, University of Minnesota
Jere R. Francis, University of Iowa
Robert W. Holthausen, University of Pennsylvania
Edmund Outslay, Michigan State University
Rick Young, Ohio State University

CORPORATE ACCOUNTING POLICY SEMINAR COMMITTEE

Charges: 1. To plan the 1993 Corporate Accounting Policy Seminar in accordance with established AAA policies.
2. To arrange to have the seminar conducted and administered.
3. To report on the results and effectiveness of the seminar and to recommend changes by November 1, 1993.

Chairperson: Roland E. Dukes, University of Washington
Alison H. Ashton, Duke University
James Don Edwards, University of Georgia
Eugene Flegm, General Motors Corporation
Barron H. Harvey, Howard University
William J. Ihlanfeldt, Shell Oil Company
Krishna G. Palepu, Harvard University
Ray G. Stephens, Kent State University

DELOITTE & TOUCHE WILDMAN MEDAL AWARD COMMITTEE

Charge: To administer the John R. Wildman Medal award program in accordance with the terms of the grant from the Deloitte & Touche Foundation.

Chairperson: Clyde P. Stickney, Dartmouth College
James R. Frederickson, Indiana University
Helen Gerrard, Miami University-Middleton
Bruce L. Oliver, Rochester Institute of Technology
Gerald Searfoss, Deloitte & Touche
Robert D. Taylor, Florida Atlantic University
Neal R. VanZante, Texas A&I University

DOCTORAL CONSORTIUM COMMITTEE

Charges: 1. To plan the 1993 Doctoral Consortium in accordance with established AAA policies.
2. To arrange to have the consortium conducted and administered.
3. To report on the results and effectiveness of the consortium and to recommend changes by August 1, 1993.

Chairperson: Abbie Smith, University of Chicago
Froystein Gjesdal, Norwegian School of Economics and Business
Betty R. Jackson, University of Colorado at Boulder
Russell J. Lundholm, Stanford University
Laureen Maines, Duke University

DOCTORAL FELLOWSHIPS COMMITTEE

Charges: 1. To select the 1992–1993 recipients of American Accounting Association fellowship grants using the criteria approved by the Executive Committee.

2. To develop a recommendation for the Executive Committee on the future funding of doctoral fellowships.

Chairperson: Robert G. May, University of Texas at Austin
Russell M. Barefield, University of Georgia
Paul Danos, University of Michigan
Ross L. Watts, University of Rochester

DOCTORAL PROGRAMS CONFERENCE COMMITTEE

Charges: 1. To finalize the plans for the 1993 Doctoral Programs Conference in accordance with established AAA policies.

2. To arrange to have the conference conducted and administered.

3. To report on the results and effectiveness of the conference and to recommend changes by June 1, 1993.

Chairperson: J. Hal Reneau, Arizona State University
Kenneth Austin, Memphis State University
Kung H. Chen, University of Nebraska
Susan Haka, Michigan State University
Douglas A. Snowball, University of Florida
Bill N. Schwartz, Virginia Commonwealth University

FINANCE COMMITTEE

Charges: 1. To prepare financial goals and strategies for the Association.

2. To evaluate additional or alternative ways to finance existing Association activities.

3. To work with the 1992–1993 President-Elect in preparing the 1993–1994 budget.

4. To recommend how AAA funds should be invested in accordance with established Executive Committee policies and procedures.

Chairperson: Jane O. Burns, Texas Tech University
Robert J. Capettini, San Diego State University
Anita Sawyer Hollander, Florida State University
Bernard J. Milano, KPMG Peat Marwick
Paula B. Thomas, Middle Tennessee State University
David A. Wilson, Ernst & Young

FINANCIAL ACCOUNTING STANDARDS COMMITTEE

Charges. 1. To evaluate selected discussion memoranda and exposure drafts related to financial accounting and reporting in the private sector as they are released by the FASB and to respond to the FASB in writing and by appearing at selected public hearings.

2. To be cognizant of emerging issues related to financial accounting and reporting in the private sector through materials issued by the Emerging Issues Task Force and other relevant sources with evaluation, where appropriate, of the significance and potential implications of such issues.

3. To meet with the FASB, normally on an annual basis.

4. To explore the publication of selected responses by the committee in *Horizons* or other appropriate outlets.

5. To consider and promote ways to increase the academic input to the standard-setting process.

6. To provide input to the Financial Reporting Issues Conference Committee on the appropriate content and structure of the 1993 conference.

Chairperson: Mary E. Barth, Harvard University
Daniel W. Collins, University of Iowa
G. Michael Crooch, Arthur Andersen & Co.
John A. Elliott, Cornell University
Thomas J. Frecka, University of Notre Dame
Eugene A. Imhoff, Jr., University of Michigan
Wayne R. Landsman, University of North Carolina
Charles L. McDonald, University of Florida
D. Gerald Searfoss, Deloitte & Touche

FINANCIAL REPORTING ISSUES CONFERENCE COMMITTEE

Charges: 1. To plan the 1993 Financial Accounting Standards Conference in accordance with established AAA policies.

2. To seek the advice of the Financial Accounting Standards Committee on the content and structure of the 1993 conference.

3. To arrange to have the conference conducted and administered.

4. To report on the results and effectiveness of the conference and to recommend changes by April 1, 1994.

Chairperson: Lawrence A. Tomassini, University of Illinois at Urbana-Champaign
Loyd C. Heath, University of Washington
L. Todd Johnson, Financial Accounting Standard Board
Donald T. Nicolaisen, Price Waterhouse
Lawrence Revsine, Northwestern University
Jerry J. Weygandt, University of Wisconsin-Madison

INNOVATION IN ACCOUNTING EDUCATION AWARD COMMITTEE

Charge: To select recipient(s) of the Outstanding Teaching/Curriculum Development Award using the criteria approved by the Executive Committee.

Chairperson: Zoe-Vonna Palmrose, University of Southern California
Clifford D. Brown, Bentley College
Robert E. Hoskin, University of Connecticut
David B. Smith, Claremont McKenna College
Beverly B. Terry, Central Piedmont Community College
David W. Wright, University of Michigan
Jean C. Wyer, Coopers & Lybrand

INTERNATIONAL FACULTY EXCHANGE COMMITTEE

Charges: 1. To select the professors to be involved in the program for 1992–1993.

2. To assist the Executive Director with announcements, arrangements, and follow-up evaluations for the respective visits.

Chairperson: Paul M. Healy, Massachusetts Institute of Technology
Peter Brownell, University of Melbourne
Nabil S. Elias, University of Manitoba
H. Perry Garsombke, University of Nebraska at Omaha
Paul L. Gerhardt, American Accounting Association
John Y. Lee, California State University-Los Angeles
Patricia O'Brien, University of Michigan
Kathleen E. Sinning, Western Michigan University
Jenice P. Stewart, University of Missouri-Columbia

MEMBERSHIP COMMITTEE

Charges: 1. To increase the quantity and quality of the membership of the Association.

2. To increase the quantity of the subscriptions to all Association journals, but especially *Accounting Horizons* and *Issues in Accounting Education.*

3. To develop a strategy for communicating benefits of AAA membership to two-year college faculty.

Co-Chairpersons: Dennis Greer, Utah Valley Community College and
James D. Stice, Brigham Young University
Don F. Arnold, Union College
George F. Aulbach, West Texas State University
Fran L. Ayres, University of Oklahoma
Paul R. Bahnson, University of Montana
William M. Baker, Appalachian State University
Thomas E. Balke, University of Nebraska
Steve Balsam, Temple University
Joyce E. Berg, Washington University
Sharon Bossung, SUNY-Geneseo
Carl W. Brewer, Sam Houston State University
Joe K. Cheung, George Mason University
John V. Crampton, Southeast Missouri State University
Catherine A. Craycraft, University of New Hampshire
Margaret F. Criner, University of Maine
Mike L. Davis, Lehigh University
Donald R. Deis, Jr., Louisiana State University
Marc DePree, University of Southern Mississippi
Mike L. Ettridge, Washington State University
Ehsan H. Feroz, University of Minnesota-Duluth
Karen A. Fortin, Georgia Southern University
Anneta Gibson, Walla Walla College
Peter R. Grierson, Slippery Rock University
Phil J. Harmelink, University of New Orleans
Ruth C. Harris, Virginia Union University
Carla K. Hayn, Northwestern University
Frederick D. Heiman, Adelphi University

John Hill, Indiana University
Vera H. Hollins, Southern University
Robert E. Holtfreter, Fort Hays State University
Anita Hope, Tarrant County Junior College
Jerry B. Hoth, California State University-Fullerton
Rudolph A. Jacob, Pace University
Teresa A. John, New York University
Dave R. Koeppen, Boise State University
Howard J. Lawrence, Christian Brothers University
Mary J. Loyland, University of North Dakota
Daphne Main, University of Maryland
M. C. Maloo, Towson State University
John B. MacArthur, University of Northern Iowa
Kathryn M. Means, Florida Atlantic University
Paul H. Mihalek, University of Hartford
Susan Moncada, Indiana State University
Alastair J. Murdoch, University of Saskatchewan
Gordian A. Ndubizu, Drexel University
Kent Poff, East Carolina University
Geraldine Powers, Northern Essex Community College
Marsha Puro, Southern Illinois at Edwardsville
Janis R. Reeder, University of Delaware
Priscilla R. Reis, Bucknell University
Terry K. Sheldahl, Savannah State College
Elaine Simpson, St. Louis Community College at Florissant Valley
Farhad Simyar, Concordia University
Hans D. Sprohge, Wright State University
E. Kay Stice, Rice University
Jane B. Stockard, Georgia College
Mary C. Tichich, The American University
Deborah H. Turner, Georgia Institute of Technology
James P. Trebby, Marquette University
Richard M. Tubbs, University of Iowa
Carol F. Venable, San Diego State University
Larry Weiss, Tulane University
David C. Yang, University of Hawaii at Manoa
Mary Ziebell, Seattle University

MINORITY FACULTY DEVELOPMENT COMMITTEE

Charges: 1. To explore additional ways in which potential minority candidates can be encouraged to apply to accounting doctoral programs and to pursue professional careers in accounting education.

2. To explore ways to encourage the development of minority faculty members.

3. To examine and monitor minority recruitment programs of other professional organizations, such as the AICPA.

Chairperson: Shirley A. Friar, Howard University
Vivian L. Carpenter, Florida A&M University
Rameshwar D. Gupta, Jackson State University
Theresa A. Hammond, Boston College
Willie E. Gist, University of Oklahoma

NEW FACULTY CONSORTIUM COMMITTEE

Charges: 1. To plan the 1993 New Faculty Consortium in accordance with established AAA policies.

2. To arrange to have the consortium conducted and administered.

3. To report on the results and effectiveness of the consortium and to recommend changes by March 1, 1993.

Chairperson: Mikel G. Tiller, Indiana University
Silvia A. Madeo, University of Missouri-St. Louis
Jane F. Mutchler, Penn State University
Dennis Reigle, Arthur Andersen & Co.
S. Mark Young, University of Colorado at Boulder

NOTABLE CONTRIBUTION TO ACCOUNTING LITERATURE AWARD SCREENING COMMITTEE

Charge: To identify works of exceptional merit from published accounting books and articles that meet the guidelines for the "Notable Contribution" awards established by the Executive Committee.

Chairperson: Frank H. Selto, University of Colorado at Boulder
Randolph P. Beatty, Southern Methodist University
Ralph L. Benke, James Madison University
Robert Bloom, John Carroll University
Germain B. Boer, Vanderbilt University
Faye A. Borthick, University of Tennessee
Michael Bromwich, London School of Economics
Larry D. Brown, SUNY-Buffalo
Larry D. Crumbley, Texas A&M University
Marie E. Dubke, Memphis State University
Margaret D. Dwyer, University of Wisconsin-Eau Claire
Leslie Eldenburg, California State University-Fresno
Jerry A. Feltham, University of British Columbia
Dale L. Flesher, University of Mississippi
Katherine C. Frazier, North Carolina State University
Fred J. Nanni, Boston University
G. Wiliam Glezen, University of Arkansas
Dave A. Guenther, University of Connecticut
Nabil M. A. H. Hassan, Wright State University
John M. Hassell, University of Texas at Arlington
Thomas Hemmer, University of Washington
Marilyn F. Johnson, University of Michigan
Alison J. Kirby, University of California-Berkeley
Lisa L. Koonce, University of Texas at Austin

Jack L. Krogstad, Creighton University
Konrad W. Kubin, Virginia Poly Institute & State University
Linda M. Malgeri, Kennesaw State College
Ella Mae Matsumura, University of Wisconsin-Madison
Tom E. McKee, East Tennessee State University
Frances E. McNair, Mississippi State University
Haim A. Mozes, Fordham University
Marilyn Neimark, CUNY-Baruch College
Jane Ou, Santa Clara University
Grover L. Porter, University of Alabama, Huntsville
Grace Pownall, Washington University
Graeme W. Rankine, Rice University
Frank R. Rayburn, University of Alabama at Birmingham
Joe G. SanMiguel, Naval Postgraduate School
Gary J. Saunders, Marshall University
Keith A. Shriver, Arizona State University
Phil H. Siegel, University of Houston-Downtown
Linda B. Specht, Trinity University
Barry H. Spicer, University of Auckland
Richard H. Tabor, Auburn University
Jim M. Wahlen, University of North Carolina
Paula A. Wilson, University of Vermont

NOTABLE CONTRIBUTION TO ACCOUNTING LITERATURE AWARD SELECTION COMMITTEE
(JOINT COMMITTEE WITH THE AICPA)

Charge: To select the recipient(s) of the award from those books and articles identified by the Screening Committee as potentially notable accounting literature contributions.

Chairperson: Rajiv D. Banker, University of Minnesota
Gerald L. Salamon, Indiana University
Anna C. Fowler, University of Texas at Austin
Robert Libby, Cornell University
Lawrence A. Ponemon, Babson College

OUTSTANDING ACCOUNTING EDUCATOR AWARD COMMITTEE

Charge: To select one or, at most, two recipients of the Outstanding Educator Award using the criteria approved by the Executive Committee.

Chairperson: Mary S. Stone, University of Alabama
Rick Antle, Yale University
Robert H. Dean, Ernst & Young
J. Richard Dietrich, University of Illinois at Urbana-Champaign
William R. Scott, University of Waterloo
Daniel T. Simon, University of Notre Dame
Gregory Waymire, Emory University

PROFESSIONAL EXAMINATIONS COMMITTEE

Charges: 1. To monitor activities of the AICPA and NASBA and evaluate the implications for the CPA examination and accounting education.

2. To identify activities the Association should consider undertaking that deal with the effect of professional examinations on accounting education.

3. To coordinate all activities and recommendations with the Accounting Education Advisory Committee.

Chairperson: Kenneth A. Smith, Idaho State University
Joyce S. Allen, Xavier University
James Blum, American Institute of CPAs
Pamela J. Erickson, Vanderbilt University
Robert Kay, New York University
John A. Marts, University of North Carolina at Wilmington
Howard R. Toole, San Diego State University

PROFESSIONALISM AND ETHICS SEMINAR COMMITTEE

Charges: 1. To plan the 1993 Professionalism and Ethics Seminar in accordance with established AAA policies.

2. To arrange to have the seminar conducted and administered.

3. To continue the development of suitable educational materials related to professionalism and ethics.

4. To report on the results and effectiveness of the seminar and to recommend changes by July 1, 1993.

Chairperson: Mary Beth Armstrong, California Polytechnic State University
W. Steve Albrecht, Brigham Young University
William H. Beaver, Stanford University
Mary S. Doucet, North Carolina A&T State University
James C. Gaa, McMaster University
Rhoda C. Icerman, Florida State University
Steven M. Mintz, Southwest Texas State University
Robert G. Ruland, Northeastern University

PROGRAM ADVISORY COMMITTEE

Charge: To assist the President and the Executive Director in developing the technical program for the 1993 Annual Meeting.

Co-Chairpersons: Victor L. Bernard, University of Michigan and
James H. Pratt, Indiana University
Robert H. Ashton, Duke University
Richard E. Baker, Northern Illinois University
Georgio Behr, University of St. Gall
Stanley F. Biggs, University of Connecticut
Frederick D. S. Choi, New York University
Robert Cooper, Harvard University
Robin Cooper, Clarements McKenna College
Jon S. Davis, University of Colorado
Mark L. DeFond, University of Southern California
Douglas V. DeJong, University of Iowa
Kenneth R. Ferris, American Graduate School of International Management
S. Michael Groomer, Indiana University
John R. M. Hand, University of Chicago
Trevor H. Harris, Columbia University
Ronald W. Hilton, Cornell University
Sally M. Jones, University of Texas at Austin
Lauren Kelly, University of Washington
James C. Kinard, Ohio State University
Richard A. Lambert, Stanford University
Baruch Lev, University of California-Berkeley
Leonard T. Long, Fisher Junior College
John O. Mason, Jr., University of Alabama
William F. Messier, University of Florida
Frederick L. Neumann, University of Illinois
D. Paul Newman, University of Texas at Austin
R. David Plumlee, Kansas State University
Judy D. Rayburn, University of Minnesota
Stephan E. Sefcik, University of Washington
Douglas A. Shackelford, University of North Carolina
Wayne H. Shaw, Cornell University
Terrence J. Shevlin, University of Washington
James F. Smith, University of Massachusetts
Mark E. Zmijewski, University of Chicago

Sections Liaison Representatives

Accounting, Behavior and Organization Section—Gary J. Mann, University of Texas at El Paso
Administrators of Accounting Programs Group—Robert G. May, University of Texas at Austin
American Taxation Association—Susan L. Nordhauser, University of Texas at San Antonio
Artificial Intelligence and Experts Systems Section—Daniel E. O'Leary, University of Southern California
Auditing—Jean Bedard, Universite of Laval
Gender Issues in Accounting (Provisional Section)—Karen L. Hooks, University of South Florida
Government and Nonprofit—Kris K. Raman, University of North Texas
Information Systems/Management Advisory Services—Martha M. Eining, University of Utah
International Accounting Section—Abdel M. Agami, Old Dominion University
Management Accounting—Cynthia D. Heagy, University of Houston-Clear Lake
Public Interest Section—Fahrettin Okcabol, Hofstra University
Two-Year College—Linda Lessing, SUNY—Farmingdale
Teaching and Curricula Section—E. Kent St. Pierre, James Madison University

PUBLICATIONS COMMITTEE

Charges: 1. To monitor all aspects of the Association's publications program.

2. To report periodically to the President and Executive Committee on all matters of concern in the field of publications.

3. To recommend candidates for editor or editor-elect positions as appropriate.

4. To recommend AAA publications policies to the Executive Committee as deemed desirable.

Chairperson: Kenneth J. Euske, Naval Postgraduate School
Jacob G. Birnberg, University of Pittsburgh
Frank Collins, University of Miami
Daniel L. Jensen, The Ohio State University
Maureen F. McNichols, Stanford University
Terrence B. O'Keefe, University of Oregon
John Willingham, KPMG Peat Marwick Main & Co.

RELATIONS WITH TWO-YEAR COLLEGE FACULTY

Charges: 1. Develop possible actions the American Accounting Association can take to increase the interaction between two-year college and four-year college faculty.

2. Because of the critical role of the introductory accounting course in attracting the best and brightest students to accounting and in providing the base of accounting knowledge on which future courses build, and because a large percentage of accounting students take introductory accounting at two-year colleges, prepare a plan for coordinating introductory accounting offerings at two- and four-year colleges.

3. Suggest ways that the American Accounting Association can provide services that are attractive to two-year college faculty.

Co-Chairpersons: Lynn Paluska, Nassau Community College and
Joseph J. Schultz, Arizona State University
Sam Asamoah, Jefferson Community College
Leonard Eugene Berry, Georgia State University
Martha Farmer, Augusta College
N. Allen Ford, University of Kansas
Blanca Gonzales, Miami Dade Community College
Charles E. Lewis, Mesa Community College

RESEARCH ADVISORY COMMITTEE

Charges: 1. To monitor the accounting research activities of other accounting associations, e.g., AICPA, FEI, IIA, and NAA.

2. To advise the Director of Research in the development and administration of the Association's program in accounting research.

3. To suggest ways to encourage research in accounting, especially applied and integrative research dealing with all accounting functions in enterprise settings.

Chairperson: Katherine Schipper, University of Chicago
Julie H. Collins, University of North Carolina
Joel S. Demski, Yale University
David F. Larcker, University of Pennsylvania
Barry L. Lewis, University of Colorado at Boulder

SECURITIES AND EXCHANGE COMMISSION LIAISON COMMITTEE

Charges: 1. To prepare and transmit a response to the staff of the Securities and Exchange Commission on proposals soliciting views from the public concerning auditing and financial accounting and reporting.

2. To make recommendations and to develop program(s) that will assist teaching of and conducting research in topics concerning the SEC and its activities.

3. To conduct activities as appropriate to assist communications and interactions between the SEC and the membership of the Association.

Chairperson: Thomas R. Weirich, Central Michigan University
James C. Flagg, Texas A&M University
Marcia S. Niles, University of Idaho
Robert W. Rouse, Auburn University
Robert J. Sack, University of Virginia
Jack E. Wilkerson, Jr., Wake Forest University

SEVENTH INTERNATIONAL CONFERENCE ON ACCOUNTING EDUCATION (1992) PLANNING COMMITTEE

Charge: To continue planning for the conference.

Chairperson: Belverd E. Needles, Jr., DePaul University
Thomas A. Gavin, University of Tennessee at Chattanooga
Sidney J. Gray, Glasgow University
W. Morley Lemon, University of Waterloo
William Markell, University of Delaware
Seigo Nakajima, Ferris Jogakuin Women's College
Walter F. O'Connor, Fordham University
Norlin G. Rueschhoff, University of Notre Dame
Murray C. Wells, University of Sydney

TOUR OF AECC GRANT SCHOOLS COMMITTEE

Charges: 1. To plan and conduct a tour of selected AECC grant schools (and posssibly others undertaking curriculum change) immediately preceding the Seventh International Conference on Accounting Education.

2. To change a fee for the tour that will make it self-supporting.

3. To prepare a report on the tour to be submitted to the Executive Committee by January 15, 1993.

Chairperson: Gary K. Meek, Oklahoma State University
Sidney J. Gray, Glasgow University
Melvin C. O'Connor, Michigan State University
Martin S. Putterill, University of Auckland
Murray C. Wells, University of Sydney

TRUEBLOOD SEMINARS COMMITTEE

Charge: To assist Deloitte & Touche in the conduct of the Trueblood Seminars.

Chairperson: James G. Manegold, University of Southern California
Robert N. West, University of Virginia
Shahrokh M. Saudagaran, Santa Clara University
E. Daniel Smith, University of Georgia

OCCASIONAL REVIEWERS AND ASSOCIATE EDITORS

The following faculty members have graciously donated their time and expertise to assist in evaluating manuscripts for *The Accounting Review* between June 1991 and November 1992. Their support is greatly appreciated. *The Editor.*

Jeffery Abarbanell, *University of Michigan*
Mohammad J. Abdolmohammadi, *Bentley College*
Margaret A. Abernethy, *The University of Melbourne*
Noel Addy, *Mississippi State University*
Giovanni Barone Adesi, *University of Alberta*
Anwer Ahmed, *University of Florida*
Bipin Ajinkya, *University of Florida*
Mimi Alciatore, *Southern Methodist University*
Michael Alles, *University of Texas at Austin*
Ashiq Ali, *Columbia University*
Amin E. Amershi, *University of Minnesota*
Eli Amir, *Columbia University*
Brenda H. Anderson, *Boston University*
Stephen Asare, *University of Florida*
Alison Ashton, *Duke University*
Anthony A. Atkinson, *Mount Allison University*
Vidya Awasthi, *Santa Clara University*
William Baber, *George Washington University*
Stanley Baiman, *University of Pennsylvania*
K. R. Balachandran, *New York University*
Michael Bamber, *University of Georgia*
Satisprasad Bandyopadhya, *University of Waterloo*
Rajiv D. Banker, *University of Minnesota*
Sasson Bar-Yosef, *The Hebrew University of Jerusalem*
Jon W. Bartley, *North Carolina State University*
Eliahu Bartov, *New York University*
Allen W. Bathke, Jr., *Florida State University*
William Beaver, *Stanford University*
Gordon Bechtel, *University of Florida*
Paul Beck, *University of Illinois*
Jean C. Bedard, *University of Connecticut*
George J. Benston, *Emory University*
Gary C. Biddle, *University of Washington*
Stanley Biggs, *University of Connecticut*
Walter Blacconiere, *University of Southern California*
Robert Bowen, *University of Washington*
Phelim P. Boyle, *University of Waterloo*
Michael E. Bradbury, *Auckland University*
Robert Bricker, *Case Western Reserve University*
Clifton Brown, *University of Illinois–Urbana-Champaign*
Lawrence Brown, *SUNY at Buffalo*
Philip R. Brown, *University of Western Australia*
Frank Buckless, *North Carolina State University*
Thomas Buchman, *University of Colorado at Boulder*
Maria L. Bullen, *Georgia State University*
Robert Bushman, *The University of Chicago*
Jeffrey L. Callen, *The Hebrew University of Jerusalem*
Paul K. Chaney, *Vanderbilt University*
John J. Cheh, *New York University*
Kevin Chen, *Rutgers University*
Peter Cheng, *Purdue University*
Joseph K. Cheung, *George Mason University*
Frederick D. S. Choi, *New York University*
Freddie Choo, *San Francisco State University*
Mary York Christ, *Pennsylvania State University*
Bryan Church, *Wilfrid Laurier University*
Gregory Clinch, *University of New South Wales*
Daniel Collins, *University of Iowa*
Jean C. Cooper, *University of Kentucky*
Robin Cooper, *Claremont McKenna College*
Mark A. Covaleski, *University of Wisconsin–Madison*
Dean Crawford, *Oakland University*
William Cready, *Texas A&M University*
Somnath Das, *University of California–Berkeley*
Larry R. Davis, *University of Minnesota*
Michael L. Davis, *Lehigh University*
Edward Deakin, *University of Texas at Austin*
Mark L. DeFond, *University of Southern California*
Donald R. Deis, *Louisiana State University*
Stephen J. Dempsey, *University of Vermont*
Rajiv Dewan, *Northwestern University*
J. Richard Dietrich, *University of Illinois*
William Dilla, *University of Illinois*
Amitabh Dugar, *Michigan State University*
Gordon Leon Duke, *University of Minnesota*
Alan S. Dunk, *University of Western Sydney*
Richard B. Dusenbury, *Florida State University*
Ronald A. Dye, *Northwestern University*
Peter D. Easton, *Macquarrie University*
Samir El-Gazzar, *Pace University*
Pieter Elgers, *University of Massachusetts*
Rachamin Elitzur, *University of Toronto*
Dan Elnathan, *University of Southern California*
Kirsten M. Ely, *University of California–Los Angeles*
Craig Emby, *Simon Fraser University*
Robert D. Emerson, *University of Florida*
Charles Enis, *Pennsylvania State University*
Brian Erard, *University of Toronto*
Michael L. Ettredge, *Washington State University*
Mahmoud Ezzamel, *Manchester Institute of Science & Technology*
Patricia M. Fairfield, *Georgetown University*
William L. Felix, Jr., *University of Arizona*
Jerry Feltham, *University of British Columbia*
Ehsan H. Feroz, *University of Minnesota–Duluth*
Lourdes D. Ferreira, *University of Baltimore*
Catherine Finger, *Illinois State University*
David Finley, *Simon Fraser University*

Joseph Fisher, *Wake Forest University*
James C. Flagg, *Texas A&M University*
Richard K. Fleischman, *John Carroll University*
George J. Foster, *Stanford University*
Taylor W. Foster, III, *New Mexico State University*
Jennifer Francis, *University of Chicago*
Robert Freeman, *University of Texas–Austin*
Carol Frost, *Washington University*
Veronique Frucot, *Syracuse University*
John C. Gardner, *SUNY–Binghamton*
Esther Gal-or, *University of Pittsburgh*
Jennifer J. Gaver, *University of Georgia*
Kenneth M. Gaver, *University of Georgia*
Michael Gift, *University of South Florida*
Gary Giroux, *Texas A&M University*
Froystein Gjesdal, *Norwegian School of Economics and Business*
G. William Glezen, *University of Arkansas*
James T. Godfrey, *George Mason University*
Joseph Godwin, *University of Georgia*
V. Gopalakrishnan, *George Mason University*
Suresh Govindaraj, *New York University*
V. Govindarajan, *Dartmouth College*
Martin F. Grace, *Southern Methodist University*
Julia Grant, *Case Western Reserve University*
Sidney Gray, *University of Glasgow*
Paul Griffin, *University of California–Davis*
Richard A. Grimlund, *University of Iowa*
David A. Guenther, *University of Connecticut*
Mahendra Gupta, *Washington University*
Karl Hackenbrack, *University of Florida*
James R. Hamill, *University of Oklahoma*
John Hand, *University of Chicago*
Stephen C. Hansen, *Stanford University*
Adrian M. Harrell, *University of South Carolina*
David Harris, *University of Michigan*
In-Mu Haw, *Texas Christian University*
Robert D. Hayes, *Tennessee State University*
Carla K. Hayn, *Northwestern University*
Richard Helleloid, *University of Wisconsin–Eau Claire*
Mark Hirst, *University of New South Wales*
Peggy Hite, *Indiana University*
Keith Houghton, *University of Melbourne*
Carol Houston, *San Diego State University*
Mohamed E. A. Hussein, *University of Connecticut*
Yuchang Hwang, *University of Pittsburgh*
Eugene Imhoff, Jr., *University of Michigan*
Raffi Indjejikian, *University of Chicago*
Robert W. Ingram, *University of Alabama*
Karim Jamal, *University of Alberta*
Christopher James, *University of Florida*
Cynthia Jeffrey, *Iowa State University*
Ross G. Jennings, *University of Texas at Austin*
Robert E. Jensen, *Trinity University*
James Jiambalvo, *University of Washington*
W. Bruce Johnson, *University of Iowa*
Frederick L. Jones, *University of Massachusetts–Dartmouth*
Scott K. Jones, *University of Delaware*
James S. Jordan, *University of Minnesota*
Woon-Oh Jung, *University of Illinois*
Sok-Hyon Kang, *Carnegie Mellon University*
Robert S. Kaplan, *Harvard University*
Steven E. Kaplan, *Arizona State University*
Duane Kennedy, *University of Waterloo*
Inder Khurana, *University of Missouri–Columbia*
Jae-Oh Kim, *The Ohio State University*
Ronald King, *Washington University*
William Kinney, Jr., *University of Texas–Austin*
Alison J. Kirby, *University of California–Berkeley*
Michael C. Knapp, *University of Oklahoma*
Gary Koehler, *University of Florida*
Leslie Kren, *University of Wisconsin–Milwaukee*
Jack L. Krogstad, *Creighton University*
William Kross, *Purdue University*
Wayne Landsman, *University of North Carolina*
William Lanen, *University of Michigan*
Amy H. Lau, *Oklahoma State University*
Raef A. Lawson, *SUNY–Albany*
Marc J. LeClere, *University of Maryland–College Park*
Charles M. C. Lee, *University of Michigan*
Robert Alan Leitch, *University of South Carolina*
John C. Lere, *St. Cloud State University*
Shu-hsing Li, *Rutgers University*
Stephen Limberg, *University of Texas–Austin*
Robert Lipe, *University of Michigan*
Gerald Lobo, *Syracuse University*
Russell Lundholm, *Stanford University*
Tom Lys, *Northwestern University*
Joseph Magliolo, III, *University of Pennsylvania*
M. Laurentius Marais, *University of Chicago*
Laureen Maines, *Duke University*
Ella Mae Matsumura, *University of Wisconsin–Madison*
William E. McCarthy, *Michigan State University*
Gary McGill, *University of Florida*
Morris McInnes, *Suffolk University*
Janet Meade, *University of Houston*
Nahum Melumad, *Stanford University*
Krishnagopal Menon, *Boston University*
Yaw M. Mensah, *Rutgers University*
Rayman D. Meservy, *Brigham Young University*
Shehzad Mian, *Emory University*
James Millar, *University of Arkansas*
H. Fred Mittelstaedt, *University of Notre Dame*
Dale Morse, *University of Oregon*
O. Doug Moses, *Naval Postgraduate School*
Susan E. Moyer, *University of Washington*
Waleed Muhanna, *The Ohio State University*
Ari Mukerji, *University of Minnesota*
Siva Nathan, *Michigan State University*
Marcia S. Niles, *University of Idaho*
Mark W. Nelson, *Cornell University*
George E. Nogler, *Suffolk University*
Patricia O'Brien, *University of Michigan*
Terrence B. O'Keefe, *University of Oregon*
Cherie O'Neill, *University of South Florida*
Thomas C. Omer, *University of Illinois*

Jane Ou, *Santa Clara University*
Tae-Young Paik, *New York University*
Kurt Pany, *Arizona State University*
Victor Pastena, *SUNY–University at Buffalo*
Evelyn R. Patterson, *University of Michigan*
Yoram Peles, *The Hebrew University of Jerusalem*
James M. Peters, *Carnegie Mellon University*
Kathy R. Petroni, *Michigan State University*
Thomas J. Phillips, Jr., *Louisiana Technological University*
Morton Pincus, *Washington University*
Larry Ponemon, *Babson College*
Grace Pownall, *Washington University*
Jamie H. Pratt, *Indiana University*
Eric G. Press, *Temple University*
Michael Prietula, *Carnegie Mellon University*
Madhav Rajan, *University of Pennsylvania*
Konduro Ramakrishnan, *Carnegie Mellon University*
Ram T. S. Ramakrishnan, *Columbia University*
Kris Raman, *University of North Texas*
K. Ramesh, *Northwestern University*
Gita Rao, *Kidder, Peabody & Co., Inc.*
Ronald Rasch, *University of Florida*
Judy Rayburn, *University of Minnesota*
Philip R. Regier, *Arizona State University*
Stefan Reichelstein, *University of California–Berkeley*
Jane L. Reimers, *Florida State University*
Sarah Reiter, *SUNY–Binghamton*
J. Hal Reneau, *Arizona State University*
Richard D. Rennie, *University of Regina*
David Ricchiute, *University of Notre Dame*
Alan J. Richardson, *Queen's University*
Walter Robbins, *University of Alabama*
Robin Roberts, *University of Missouri–Columbia*
Waymond Rodgers, *University of California–Irvine*
Kermit J. Rohrbach, *University of Oklahoma*
Uri Ronnen, *Yale University*
James D. Rosenfeld, *Emory University*
Alan Roshwalb, *Georgetown University*
Marc A. Rubin, *Miami University*
William Salatka, *University of Kansas*
Richard Sansing, *Yale University*
David Sappington, *University of Florida*
Bharat Sarath, *New York Unversity*
Shahrokh M. Saudagaran, *Santa Clara University*
Michael Schadewald, *University of Wisconsin–Milwaukee*
Thomas Schaefer, *Florida State University*
David Schkade, *University of Texas at Austin*
Arnold Schneider, *Georgia Institute of Technology*
Douglas A. Schroeder, *The Ohio State University*
Hein Schreuder, *University of Limburg*
Thomas Scott, *University of Alberta*
William R. Scott, *University of Waterloo*
Frank Selto, *University of Colorado–Boulder*
P. K. Sen, *University of California–Berkeley*
Timothy L. Shaftel, *University of Kansas*
Wayne H. Shaw, *University of Colorado–Boulder*
Mohamed Shehata, *McMaster University*
Toshi Shibano, *University of California–Berkeley*
Douglas Shin, *University of Pittsburgh*
Randolph A. Shockley, *University of Georgia*
Donna J. Shores, *University of Washington*
Daniel Simon, *Indiana University Northwest*
Stephen Slutsky, *University of Florida*
Wally Smieliauskas, *University of Toronto*
Naomi S. Soderstrom, *University of Washington*
Lisa Gilbert Soo, *Boston College*
Theodore Sougiannis, *University of Illinois*
Eric Spires, *The Ohio State University*
Sri Sridharan, *Northwestern University*
Bin Srinidhi, *Rutgers University*
Michael T. Stein, *Rutgers University*
Paul Steinbart, *Memphis State University*
Ray G. Stephens, *Kent State University*
Earl K. Stice, *Rice University*
Scott E. Stickel, *University of Pennsylvania*
Christopher Stinson, *University of Washington*
Thomas Stober, *Indiana University*
Don J. Stokes, *University of Queensland*
Yoon S. Suh, *University of California–Los Angeles*
Edward P. Swanson, *Texas A&M University*
Amy P. Sweeney, *Harvard University*
Richard H. Tabor, *Auburn University*
Walter Teets, *University of Illinois*
Siew Hong Teoh, *University of California–Los Angeles*
Sundaraman Thiagarajan, *Northwestern University*
Lynda Thoman, *Purdue University*
Jacob Thomas, *Columbia University*
Robert Thompson, III, *University of Maryland*
Mark Trombley, *University of Arizona*
Brett Trueman, *University of California–Berkeley*
Richard Tubbs, *University of Iowa*
Robert Tucker, *University of Illinois at Chicago*
Richard A. Turpen, *University of Tennessee*
Carol Venable, *San Diego State University*
P. C. Venkatesh, *University of Missouri at St. Louis*
Robert E. Verrecchia, *University of Pennsylvania*
Robert L. Vigeland, *Texas Christian University*
Jayaraman Vijayaumar, *Georgia State University*
James Waegelein, *University of Kansas*
Martin Walker, *University of Manchester*
David Wallin, *The Ohio State University*
Shiing-wu Wang, *University of Arizona*
Terry D. Warfield, *University of Wisconsin–Madison*
John S. Watts, *Purdue University*
Ross L. Watts, *University of Rochester*
Joseph Weintrop, *SUNY at Buffalo*
Greg P. Whittred, *University of New South Wales*
David Williams, *The Ohio State University*
Arnold Wright, *University of New South Wales*
David Wright, *University of Michigan*
Bing, Xiang, *University of Calgary*
Jai-Zheng Xie, *University of Alberta*
James Yardley, *Virginia Polytechnic and State University*

Jeffrey Yost, *University of Florida*
S. Mark Young, *University of Southern California*
Paul Zarowin, *New York University*
Christine Zavgren, *Clarkson University*
David Ziebart, *University of Illinois*
Amir Ziv, *Yale University*

THE ACCOUNTING REVIEW
Vol. 68, No. 1
January 1993
pp. 205–212

BOOK REVIEWS

Shane R. Moriarity, Editor

Editor's Note: Two copies of books for review should be sent to Professor Shane Moriarity, School of Accounting, University of Oklahoma, Norman, OK 73019. The policy of the *Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

R. J. CHAMBERS, *Foundations of Accounting* (Geelong, Australia: Deakin University Press, 1991, pp. v, 216, A$38.00, paper).

Professor Chambers continues his advocacy of exit prices in this book. His objective is to use the ideas from scholarly disciplines other than accounting as the foundation for his arguments for continuously contemporary accounting and against historical cost and replacement cost accounting models. Accordingly, the book contains eleven sections with titles such as Commercial foundations, Legal foundations, and Financial foundations. If, for example, we examine the section on Legal foundations, we find that one of his central themes is that legal form represents the substance of accounting events. He interprets this idea as supporting the measurement of liabilities at the legal amount for which they can be settled. This type of approach is representative of the entire book and he argues that all of the ideas come together to provide a logical foundation undermining conventional accounting and supporting continuously contemporary accounting. A more appropriate title for the book would be *A Foundation for a Theory of Accounting*.

An unusual, if not unique, device in the book is a series of questions integrated into the text. Answers to the questions are provided in the last section of the book. The questions and answers are designed to provide additional support for the accounting model that is proposed. In the end, however, the book is another exposition of prescriptive accounting theory in the traditional style.

A key question about the book that is unanswered in this review is the degree to which it adds to Professor Chambers' previous work. I am not familiar enough with his work to make this determination. However, my guess is that most of the important ideas in this book are embedded in his earlier published work. The book is so carefully put together, the series of questions so cleverly constructed, and the sources referred to so broad and varied that it is obvious that the book is the result of many years of scholarly thought. On the other hand, there are no obvious new ideas or theories relied upon that would have prevented the book from having been published in 1970 or, perhaps, even in 1960. Only four post-1970 nonaccounting references are listed in the fourteen pages of references. If we can borrow from the language of literary scholars, it is appropriate to say that few, if any, of the accounting references are currently "in the canon" of academic accounting research.

Professor Chambers has explicitly eliminated undergraduate students as potential readers of the book since he presumes "a good working knowledge of the state of the art, as expounded in the standard textbooks and educational programs" (p. 1). Be warned that the book is not a quick and easy read even though it is relatively short and Professor Chambers' style is concise and nonpedantic. Therefore, scholars merely curious about prescriptive accounting theory probably will not find reading it worthwhile. Scholars no longer interested in traditional prescriptive theory will find nothing in the book to interest them. Those scholars with a serious interest in traditional prescriptive theory should read the book. Then they can make the judgment about what is appropriate for their graduate students.

DALE A. BUCKMASTER
Professor of Accounting
University of Delaware

HOWARD R. DAVIA, PATRICK C. COGGINS, JOHN C. WIDEMAN and JOSEPH T. KASTANTIN, *Management Accountant's Guide to Fraud Discovery and Control* (New York: John Wiley & Sons, Inc., 1992, pp. xvii, 213, $55.00).

This book presents a fairly comprehensive description of the environment of fraud and fraud auditing and proposes the organization of a new association for fraud auditors. It is written in relatively simple terms (the authors purposely avoided highly technical terms and details) in order to encourage wider distribution and increase its readership among top management in business. It is hoped that top management will gain an appreciation for, and see the need to support, fraud auditing.

The authors make the point that external or financial auditors do not fully address the issue of fraud. In fact the new standard auditor's report is silent about fraud, irregularities, and illegal acts. However, these authors also cite the U.S. Chamber of Commerce estimates that employee and management fraud exceeds $100 billion annually and make a strong case for the need of fraud auditing.

An interesting proposal by the authors is to form a new association, "Fraud Auditors and Investigators Association." This association would apparently be patterned after the public accounting profession and two suggested names for its members would be "frauditors" or "fraudigators." Seven general standards for fraud auditing and investigating are proposed and field standards will be prepared and issued as research and time permits.

There are a few statements that tend to reduce the book's credibility by overstating the authors' position or making unsubstantiated statements. The following quote is an illustration.

> In essence, many perpetrators live very well off victims who never even suspect that they are being drained of resources. It is not unreasonable to assume that 2 percent of a business's or institution's gross revenues are being lost to fraud, and that 2 percent of its employees are involved in fraud. As if they are playing a slot machine that requires no coins, they keep pulling the handle and it keeps paying off, year after year. For many businesses, profits are sufficiently high that they do not miss the 2 percent (p. 64).

Overall, however, the chapter materials are well organized and describe the nature of fraud and fraud auditing very well. In addition to this description of fraud and fraud auditing, probably the best educational materials are contained in the eight cases found in Chapter 8. These cases along with their suggested solutions add a convincing sense of reality to the rest of the materials presented. On the other hand, probably the least educational materials are contained in Chapter 11, Rules of Evidence. This chapter proved to be more difficult to understand and relate to the rest of the book. Examples of how to go about gathering and using evidence would considerably improve the usefulness of this chapter.

The overall impact of the book is to sound an alarm that fraud is happening and some actions can and should be taken to protect the assets under one's responsibility. Having had significant responsibility as an internal auditor in a company with over $22 billion in sales, my experiences verify much of the materials written by the authors and I share their concern about custodial fraud where pension, welfare, and similar funds are held in trust.

I strongly recommend this book for those responsible for the management or safekeeping of assets and I believe policy makers in the insurance, pension, banking, and similar businesses would find it interesting and useful. Because this book has less than 200 pages of text and is relatively easy to read, it could be adapted into the auditing curriculum fairly easily. It would make an excellent topic for discussion in a graduate seminar and should be a "must read book" for internal auditing courses. I recommend all auditing educators read it to see how it may be integrated into their auditing programs.

JAMES W. BRACKNER
Associate Professor of Accounting
Utah State University

JOHN DONALDSON, Business Ethics: *A European Casebook* (London: Academic Press Limited, 1992 pp. xxvi, 293, £14.95, paper).

The author of this book sets out to add positive, constructive examples of European-based ethics cases to the comparatively few ethics cases existing in the business literature. Maintaining "that the key issues of business ethics are more to do with *how* rules are made, than with *what* rules are made" (ix), the author takes care to elucidate a structure for ethical analysis in the first seven chapters, and includes a glossary of ethics terms (which appears, somewhat imposingly, at the front of the book). The first chapter contains twenty short cases or examples, usually no more than a paragraph or two in length, which the author uses to illustrate various aspects of his preferred terminology and framework for ethical analysis. The first seven chapters of the book essentially present and analyze 30 cases which set forth the author's perspective on the study of ethics in general. A total of 48 cases, most of which relate to Britain and other members of the European Community, occupy approximately 100 pages of the book, a little better than a third of the total page count. For this reason, the volume's title of "Casebook," which neither refers to the author's lengthy discussion of ethics in general nor to the extent of the author's case analysis may be somewhat misleading.

The nineteen longer cases (1-10 pages or more) examine a gamut of current issues or topics on the European scene such as corporate and professional codes of ethics, environmental concerns, wage and price control policies, economic theory and policies as applied to markets, and the mass media's propagation of values. These cases customarily look at the issues from multiple perspectives, drawing information from corporate sources or other inside sources, as well as from various published reports. As the topics listed above indicate, the cases by and large address ethical issues from a macro view—that of the corporation, profession, industry, or the European community as a whole.

Almost all of the cases were chosen to illustrate so-called "normal" values in business, which vary among companies, industries, and countries for many reasons that are detailed in the text. However, only some of the cases drawn from the financial services industry exemplify "normal" values, while others illustrate lessons to be learned from scandals. In a method of case writing unusual to the reviewer, throughout the book the author includes his commentary and analysis in the case itself, as well as before or after it.

The financial services cases (3 long, 6 short) are the only ones which strike the reviewer as being directly germane to the accounting classroom. Accounting professors who seek a casebook useful for integrating accounting or control-related ethical issues into their courses would best look elsewhere.

The book lends itself more toward use as a supplementary reading for an international business or business ethics course; as such, it gives a valuable perspective on some of the current ethical concerns in Europe, as well as providing and illustrating tools for ethical case analysis at the macro level. Because of these characteristics, it should also be a useful resource for those interested in doing research in the area.

As a final note, a professor who assigns the book for class reading should be aware that the text contains a number of non-ethics-related "British English" terms, without explanation, with which U.S. students may not be familiar. An instructor-prepared list of British-American equivalent terms may enhance American students' comprehension of the text.

MARY GREENAWALT
Associate Professor of Business Administration
The Citadel

LIN FITZGERALD, ROBERT JOHNSTON, STAN BRIGNALL, RHIAN SILVESTRO and CHRISTOPHER VOSS, *Performance Measurement in Service Businesses* (London: The Chartered Institute of Management Accountants, 1991, pp. xi, 126).

This book is largely based on a two-year research project sponsored by The Chartered Institute of Management Accountants, London, England during the period of 1987 through 1989. The inter-disciplinary research team used in-depth interviews in gathering data for the project. The aim of the project was to document the state of the art in current practice in the United Kingdom in performance measurement and control in services.

The authors identify three major categories of service types: professional services, service shops, and mass services. These service types need a range of performance measurements linked to their competitive strategy. The authors establish six generic performance measurements, two which measure the results of competitive performance, and four which measure factors that determine competitive success. The two which measure the results of competitive performance are: competitiveness (e.g., sales growth and market share) and financial performance (profitability, liquidity, capital structure, and market ratios). The four which measure the determining factors are: quality of service, flexibility, resource utilization, and innovation.

The discussion of the characteristics of services: intangibility, heterogeneity, simultaneity, and perishability, explain the difficulty in measuring performance of the service sector. All of these four characteristics make it more difficult in services than in manufactured products to measure performance. The establishment of goals and the categorization of service types and measurements by the authors help the reader to better understand the otherwise complex problem of analyzing the service industry.

The relevant measures of competitive and financial performance embrace both competitor-based and customer-focused approaches. The authors find no good reasons that competitiveness and three of the four traditional financial ratios: liquidity, capital structure, and market ratios, should vary across the three generic service types. The performance measurement of profitability ratios will vary across the three service types because of their cost structures. Cost traceability affects the extent to which costs can be used for decision-making in the organization. The authors observe that cost traceability is highest in the professional services, lower in service shops, and lowest in mass services.

The book proceeds to describe, each in a separate chapter, the four factors which are involved in the determination of performance measurement. The last chapter advocates the development of an integrated system to measure performance in services, one that belongs to the whole management team, not just to management accountants.

The book represents a succinct and cogent effort to report the results of an empirical study in the service sector. The authors establish categories which help to identify the differences in types of services and to isolate the factors which are involved in the performance measurement process. It is well organized and establishes a logical and practical pattern of anlaysis. The service sector has long been neglected in cost analysis because of its intangibility and its heterogeneity, yet it may be the fastest growing sector in economically mature nations. The authors have provided a framework which can be used by scholars to pursue further research in the service sector and by practitioners to guide them in measuring the performance of their service function.

RONALD J. LEWIS
Professor Emeritus

Central Michigan University

THOMAS W. LAUER, EILEEN PEACOCK, and ARTHUR C. GRAESSER, Editors, *Questions and Information Systems* (Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 1992, pp. vii, 374, $34.50 paper, $69.95 cloth).

This book presents a collection of papers which explore the relationships between questions and information systems. All aspects of questions, including their formulation, use and impact, are explored in studying this relationship. The papers were originally presented at a 1990 multidisciplinary conference on "Questioning and Information Systems" and were subsequently revised for publication. Lauer and Graesser, in the introductory chapter, point out the importance of explicitly identifying the relationships between questions and information systems. They argue that the dynamic nature of information systems makes this relationship especially pertinent.

The diversity of the interrelationship of questions with information systems is addressed by the collection and organization of the different papers. The articles in the book were written by both academicians and individuals from industry. In addition, the backgrounds of the contributors vary from areas in psychology to artificial intelligence to accounting. The range of articles includes theory building, experiments, reviews and descriptive pieces. To assist the reader, the editors provide both a subject index and an index of authors referenced in the articles.

Questions are typically used during the development of information systems for the acquisition of knowledge and the systems analysis. Questions also impact the development of user interfaces and the use of these interfaces, particularly in querying large databases. An examination of how questions facilitate the understanding of complex decision processes is also useful in studying the interrelationships of questioning and information systems.

The articles dealing with the development of information systems lean heavily towards expert systems, but do include a nice discussion of systems analysis. While I found the articles on expert systems to be of interest, this section would have benefitted from a broader coverage of topics. These articles did provide a nice combination of theoretical and descriptive work. The area of user interface was well covered and diverse. In addition, the articles covering the topic of querying large databases provided some unique perspectives. The articles dealing with the use of questions to study complex decision processes were an interesting mixture. The issue is explored through experimental analysis, descriptive work, and theoretical development.

On the whole, the articles in this book are consistently well written and informative. The strength of the book lies in the breadth of the background of the contributors and the variety of their contributions. This book should prove interesting to those who view accounting as the dissemination of information, including the many facets related to providing that information. While this book was not written as a textbook, I think it could be used for a graduate level systems class or even as a background book for a Ph.D. seminar.

MARTHA M. EINING
Associate Professor of Accounting
University of Utah

MARCELL SCHWEITZER, ERNST TROSSMAN, and GERALD H. LAWSON, *Break-Even Analysis: Basic Model, Variants, Extensions* (Chichester: John Wiley & Sons Ltd., 1992, pp. xvi, 306, $84.95).

This book "comprehensively, and synoptically, examines the familiar variants of break-even analysis" (xi). This is underscored by four pages of symbols and abbreviations that follow the preface. If a reader is interested in a very detailed and quantitative approach to reviewing the break-even model and its variants and extensions, this book provides such a review; it is complete and very detailed. If a reader is more interested in specific applications, this is a good reference book that provides complete examples for the various portions of the model examined. However, its greatest strength (complete, detailed, comprehensive) is also its greatest weakness: the authors tend to rely too much on having to develop each piece of the puzzle using mathematical notation and in being sometimes tedious in leading the reader through the material.

The book carries a 1992 copyright, but the original German version was published in 1986 which means references are limited by the earlier publication date. There are two main parts: Chapters 2 and 3 cover the basic model and some simple variants of the model. Chapter 4 (a whopping 190 pages) covers extensions. Both Chapters 1 and 5 (introduction and summary) are quite brief (two and four pages, respectively).

The groundwork for the extensions, the major thrust of the book, is set in the first three chapters. All the basics of cost-volume-profit and break-even (BE) analysis are explored using several consistent series of examples that are drawn from German companies. While the descriptions and explanations of, say, topics in basic sensitivity analysis are overly long, the examples allow the reader to better understand and integrate the flow of mathematical notation.

Chapter 2 develops the basic BE model as well as a complete discussion of its assumptions and limitations. Chapter 3 deals with variants of BE analysis. A substantial part of this chapter is basic sensitivity analysis. There is a section on how to apply BE to replacement investments that seems muddled, however, since it deals with profits instead of cash flows and does not include the time value of money. The authors miss a chance here to incorporate discounted BE analysis. In addition, there is a long section dealing with sales risk, which does not add substantively to the original 1964 work of Jaedicke and Robichek.

The bulk of the book is in Chapter 4 dealing with extensions. The topics in this chapter are good, but it should have been divided into several chapters each dealing with a focused section of the material as is illustrated by the following. At the beginning of the chapter, the authors provide a guide to the organization of the extensions: single-stage multiproduct manufacture, multistage mutiproduct manufacture, multidimensional production functions, multiperiod BE analysis, non-linear analysis, stochastic analysis, and multiple objectives. These topics include simple subjects like weighted-average contribution margin, contribution margin per scarce resource, and linear programming. While these sections are straightforward, as with much of the other material in the book, in an effort to be complete, the authors err on the side of being tedious and too detailed with notations, tables, and graphs.

The material on multistage multiproduct manufacture is an important section since it goes well beyond what most academics and practitioners view as part of the BE model and incorporates transfer prices as intermediate goods move from one stage to the next. There is also a good section showing how the application of such a complex analysis is linked, unfortunately, to the existing cost system. This idea ties in well with the current literature dealing with revisions of how costs are accumulated and assigned (e.g., activity-based costing). The sections on subjects like multistage multiproduct manufacture and multidimensional production functions will be appealing to those who are interested in quantitative modelling. However, the authors do provide examples in each area to show that there is the possibility of actual implementation of the models.

This book is a welcome addition to a university library as a reference book. Its quantitative style and somewhat lengthy exposition make it an unlikely candidate to be used as a text, *per se*, in anything other than an advanced operations management and research class for doctoral students.

MAURICE L. HIRSCH, JR.
Professor of Accounting
Southern Illinois University at Edwardsville

ANTHONY STEELE, *Audit Risk and Audit Evidence: The Bayesian Approach to Statistical Auditing* (London: Academic Press, 1992, pp. xii, 200, $32.00).

This monograph is an outgrowth of a short series of lectures delivered by the author at the University of Lancaster in 1984–1985. After briefly examining the role of auditing in society, the author focuses on risk sharing and risk management and the need for coherent audit judgments. This background analysis is then used to motivate a discussion of a Bayesian subjective probability model for attribute-oriented audit judgments. While not specifically subdivided into two parts, this more introductory material constitutes, approximately, the first half of the monograph.

Bayesian procedures for variable sampling are addressed in the second half of the monograph. This includes (in part) a brief examination of the role of materiality, loss functions, optimal sample sizes and prior to posterior analyses. This is followed by a more extensive examination of two Bayesian models from the auditing research literature (William L. Felix, Jr., and Richard A. Grimlund, *Journal of Accounting Research*, Vol. 15 Spring 1977, pp. 23–44 and John H. McCray, *The Accounting Review*, Vol. 54 No. 1, p. 35).

Other than noting that the students attending his lecture series needed a text, the author is silent with repect to how the monograph might be utilized in a pedagogical setting. The author does express the belief "that auditors need to study formal models of judgment in order that they understand the informal models which underlie contemporary practice..." (p. 35). However, the linkage between the two types of understanding is not subsequently developed.

Nonetheless, the author has provided an excellent primer for students, academics, and practitioners interested in learning about Bayesian statistical auditing approaches to variable sampling. The more introductory material in the first half of the monograph is also very well done. There are numerous figures, tables, numerical examples and references to related literature throughout the monograph. As an added plus, the first half of the monograph uses many U.K. sources that American readers may not have previously noted.

An important feature of the monograph is its emphasis on implementation issues. One first sees this on a more conceptual level in the care that has been taken in placing statistical sampling within the context of managing audit risk. Subsequently, one finds detailed discussions, exhibits, computer code and availability information for a variety of supporting computer software. This material focuses on equivalent prior samples, computer assisted data analysis (CADA), using beta distributions, selecting materiality (AUDITPLANNER), sample size planning, several general purpose Bayesian procedures, and so on.

A major aspect of the monograph is the coverage given to the Felix and Grimlund (1977) beta-normal Bayesian model and the McCray (1984) quasi-Bayesian model. The emphasis of these two sections of the monograph is on numerical examples and how these models might be used as practical auditing procedures. Nevertheless, these are somewhat mathematical demanding sections of the monograph which require, for a complete understanding, some familiarity with integral calculus and likelihood functions. While the author has taken many steps "to reduce the mathematics to its simplest" (p. xi), readers of this and related material will need a dedication to mathematical details. A dedication that, in this reviewer's experience, is not commonly observed among stateside undergraduate and master's level accounting students.

While there are many sections of the monograph that would be suitable and quite helpful for such students, given its cost ($32.00), it probably would not be justifiable to ask such students to purchase the monograph. In contrast, the monograph could be an important part of a Ph.D. seminar that wishes to take a detailed look at modern Bayesian approach to audit sampling. Further, the monograph will be very helpful to academics and practitioners with a special interest in audit sampling.

RICHARD A. GRIMLUND
Associate Professor of Accounting

University of Iowa

BOOKS RECEIVED MAY 1, 1992 TO AUGUST 1, 1992

EHIEL ASH and ROBERT STRITTMATTER, *Accounting in the Soviet Union* (New York: Praeger Publishers, 1992, pp. xii, 191, $45.00).

D. M. BARNES, *Value-For-Money Audit Evidence* (Toronto: The Canadian Institute of Chartered Accountants, 1992, pp. xii, 117, C$35.00, paper).

VINOD B. BAVISHI, *Global Company Handbook: An Analysis of the Financial Performance of the World's Leading 7,500 Companies, 1992 Edition* (Princeton: Center for International Financial Analysis & Research, Inc., 1992, pp. 2,500 in 2 volumes, $395.00, paper).

PETER BOYS and JOHN FREEAR, Editors, *Accounting History 1976–1986: An Anthology* (New York: Garland Publishing, Inc., 1992, pp. xii, 402, $87.00).

T. E. COOKE and M. KIKUYA, *Financial Reporting in Japan: Regulation, Practice and Environment* (oxford: Blackwell Publishers, Ltd., 1992, pp. xii, 356, $89.95.

HUGH M. COOMBS and J. R. EDWARDS, Editors, *Local Authority Accounting Methods: Problems and Solutions, 1909–1934* (New York: Garland Publishing, Inc., 1992, pp. xiii, 369, $70.00).

GRAEME W. DEAN, MARC P. JOYE, and PAUL J. BLAYNEY, *Strategic Management Accounting Survey: Overhead Cost Allocation & Performance Evaluation Practices of Australian Manufacturers* (Sydney: University of Sydney, 1991, pp. x, 144, A$25.00, paper).

DOUGLAS T. HICKS, *Activity-Based Costing for Small and Mid-Sized Businesses: An Implementation Guide* (New York: John Wiley & Sons, 1992, pp. xii, 312, $59.95).

MARTIN N. HOOGENDOORN and LILLIANE B. BIJL, Editors, *Financial Reporting in the Nineties: Regulation and Innovation at a Crossroads* (Amsterdam: Foundation for Annual Reporting, 1992, pp. 80, g47.50, paper).

ROWAN JONES, *The History of the Financial Control Function of Local Government Accounting in the United Kingdom* (New York: Garland Publishing, Inc., 1992, pp. iv, 184, $62.00).

SIMON M. KEANE, *A Survey of the Valuation Practices of Professional Accounting Firms* (Edinburgh: The Institute of Chartered Accountants of Scotland, 1992, pp. vi, 27, £10.00).

G. H. LAWSON, Editor, *Studies in Cash Flow Accounting and Analysis: Aspects of the Interface Between Managerial Planning, Reporting and Control and External Performance Measurement* (New York: Garland Publishing, Inc., 1992, pp. v, 315, $63.00).

JOHN LESLIE LIVINGSTONE, *The Portable MBA in Finance and Accounting* (New York: John Wiley & Sons, Inc., pp. xi, 524, $24.95).

LEONARD LORENSEN, *Accounting for Liabilities* (New York: American Institute of Certified Public Accountants, 1992, pp. xv, 173, $30.00, paper).

JOHN MCMILLAN, *Games, Strategies, & Managers* (New York: Oxford University Press, 1992, pp. x, 252, $22.95).

YASUHIRO MONDEN, *Cost Management in the New Manufacturing Age: Innovations in the Japanese Automotive Industry* (Cambridge, MA: Productivity Press, 1992, pp. xxii, 176, $45.00).

RICHARD G. NEWMAN, *Supplier Price Analysis: A Guide for Purchasing, Accounting, and Financial Analysts* (New York: Quorum Books, 1992, pp. xii, 186, $47.95).

ASHEQ RAZAUR RAHMAN, *The Australian Accounting Standards Review Board: The Establishment of its Participative Review Process* (New York: Garland Publishing, Inc., 1992, pp. xxiii, 540, $73.00).

HAL F. ROSENBLUTH, *The Customer Comes Second and Other Secrets of Exceptional Service* (New York: William Morrow and Company, Inc., 1992, pp. 240, $20.00).

LIONEL SALEM, FREDERIC TESTARD, and CORALIE SALEM, *The Most Beautiful Mathematical Formulas* (New York: John Wiley & Sons, Inc., 1992, pp. xiii. 141, $19.95).

MARION G. SOBOL, GAIL E. FARRELLY, and JESSICA S. TAPER, *Shaping the Corporate Image: An Analytical Guide for Executive Decision Makers* (New York: Quorum Books, 1992, pp. xiv, 168, $39.95).

ANTHONY STEELE, *Audit Risk and Audit Evidence: The Bayesian Approach to Statistical Auditing* (London: Academic Press, 1992, pp. xii, 200, $32.00, paper).

G. D. TRITES, et al., *Information to be Included in the Annual Report to Shareholders* (Toronto: The Canadian Institute of Chartered Accountants, 1992, pp. xv, 298, C$39.50, paper).

IAN WALKER, *Buying a Company in Trouble* (Hants, England: Gower Publishing Company Limited, 1992, pp. x, 134, $49.95).

TANG YUN WEI, LYNNE CHOW, and BARRY J. COOPER, *Accounting and Finance in China: A Review of Current Practice* (Hong Kong: Longman Group (Far East) Ltd., 1992, pp. x, 224, $55.00).

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal.... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of the *Chicago Manual of Style* (13th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's International Dictionary*.

FORMAT

1. All manuscripts should be typed on one side of 8½ × 11″ good quality paper and be double spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To assure anonymous review, authors should not identify themselves directly or indirectly in their papers. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title, and affiliation, any acknowledgements, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: *All* pages, including tables, appendices, and references, should be serially numbered. The first section of the paper should be untitled and unnumbered. Major sections may be numbered in roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights, and measures. For example: *three days; 3 kilometers; 30 years.* All other numbers are expressed numerically. Generally when using approximate terms spell out the number, for example, *approximately thirty years.*

Percentages and Decimal Fractions: In nontechnical copy use the word *percent* in the text; in technical copy the symbol % is used. (See the *Chicago Manual* for discussion of these usages.)

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: a *well-presented analysis; re-form.* See *Webster's* for correct usage.

Key Words: The synopsis is to be followed by four *key words* that will assist in indexing the paper.

SYNOPSIS

A synopsis of about 400 words should be presented on a separate page immediately preceding the text. The synopsis also serves as the "introduction." The synopsis should be nonmathematical and include a readable summary of the research question, method, and the significance of the findings and contribution. The title, but not the author's name or other identification designations, should appear on the synopsis page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of t text. Each should bear an arabic number and a complete title indicating the exact contents of the table figure.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (s below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: f example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 19 115).
3. When the reference list contains more than one work of an author published in the same year, the suf *a*, *b*, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jon (1987, 115) says. . . ."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (A ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutio work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises, or court cases, citations acceptable in law revie should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Ea entry should contain all data necessary for unambiguous identification. With the author-date system, use t following format recommended by the *Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the insti tion responsible for the citation.
2. Use author's initials instead of proper names.
3. Dates of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) should be listed in chronological order of publication. Two or mo works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Repor 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2d ed. Homewood, IL: D Jones-Irwin.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northweste University, Evanston, IL.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Ca bridge, United Kingdom: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–36.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the Am ican Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. disser tion, University of Texas, Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harva Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be double spaced and numbered consecutively throughout the manuscript with superscript Arabic numerals. Footnotes are placed at the end of the text.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or other publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan, or the like) should be submitted.
3. Four copies should be submitted together with a check in U.S. funds for $50.00 for members or $100.00 for nonmembers of the AAA made payable to the American Accounting Association. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the Editor for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting.

"An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional, research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

SHIP, MAN-
ATION (Re-

Accounting Re-

24, 1992.
erly.
d *annually*: 4.
Price: $50.00.
of publication
ZIP code) (not
rasota, Florida

rters of general
rs (not printers):
Florida 34233-

addresses of
editor.
): American Ac-
Bessie Drive,

l-khalik, Fisher
of Business Ad-
ida, Gainesville,

d address): Paul
arasota, Florida

corporation, its
stated and also
names and ad-
ng or holding 1
If not owned by a
resses of the indi-
owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)

Name: American Accounting Association (A nonprofit educational association—no stockholders or "owners").

Address: 5717 Bessie Drive, Sarasota, Florida 34233-2399.

8. *Known bondholders, mortgagees, and other security holders owning or holding 1 percent or more of total amount of bonds, mortgages or other securities (If there are none, so state)*: none.

9. *For completion by nonprofit organizations authorized to mail at special rates (Section 423.12 DMM only). The purpose, function and nonprofit status of this organization and the exempt status for federal income tax purposes:* Have not changed during preceding 12 months.

10. *Extent and nature of circulation:*

A. *Total no. copies printed (Net Press Run).* Average number copies each issue during preceding 12 months: 16,000. Actual number copies of single issue published nearest to filing date: 16,000.

B. *Paid Circulation:*

1. Sales through dealers and carriers, street vendors and counter sales. Average number copies each issue during preceding 12 months: none. Actual number copies of single issue published nearest to filing date: none.

2. *Mail subscriptions.* Average number copies each issue during preceding 12 months: 13,154. Actual number copies of single issue published nearest to filing date: 13,475.

C. *Total paid circulation (sum of 10B1 and 10B2).* Average number copies each
during preceding 12 months: 13,154.
number copies of single issue published n
to filing date: 13,475.

D. *Free distribution by mail, carrier or*
means, samples, complimentary, and othe
copies. Average number copies each
during preceding 12 months: 37. Actual nu
copies of single issue published nearest to
date: 37.

E. *Total distribution (Sum of C an*
Average number copies each issue durin
ceding 12 months: 13,191. Actual n
copies of single issue published nearest to
date: 13,512.

F. *Copies not distributed:*

1. *Office use, left over, unaccounted, s*
after printing. Average number copies
issue during preceding 12 months: 2,809.
number copies of single issue published n
to filing date: 2,488.

2. *Returns from news agents.* Av
number copies each issue during precedi
months: none. Actual number copies of
issue published nearest to filing date: nor

G. *Total (Sum of E, F1 and 2—should*
net press run shown in A). Average n
copies each issue during preceding 12 m
16,000. Actual number copies of single
published nearest to filing date: 16,000.

H. *I certify that the statements made above are correct and complete.*

Paul L. Gerhardt, Executive Director

A FORUM ON

THE EFFECTS OF VIOLATING DEBT COVENANTS

1. Kevin C. W. Chen and K. C. John Wei
   Creditors' Decisions to Waive Violations of Accounting-Based Debt Covenants
2. Messod D. Beneish and Eric Press
   Costs of Technical Violation of Accounting-Based Covenants
3. Samir M. El-Gazzar
   Stock Market Effects of the Closeness to Debt Covenant Restrictions Resulting from Capitalization of Leases
4. Mary Beth Mohrman
   Debt Contracts and FAS No. 19: A Test of the Debt Covenant Hypothesis
5. Clifford W. Smith, Jr.
   A Perspective on Accounting-Based Debt Covenant Violations

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 218–232

Creditors' Decisions to Waive Violations of Accounting-Based Debt Covenants

Kevin C. W. Chen
Rutgers University at New Brunswick
K. C. John Wei
The Hong Kong University of Science and Technology

SYNOPSIS AND INTRODUCTION: Positive theory hypothesizes that accounting-based debt covenants are important factors in accounting choices. According to Watts and Zimmerman's (1990) survey, this hypothesis has generally been supported by earlier studies. That is, the closer the firm is to violating accounting covenants, the more likely managers would choose income-increasing methods.

Recently, research attention has shifted to the event of covenant violation itself. For example, Beneish and Press (1993) estimate debtors' costs of violations. Further, DeFond and Jiambalvo (1991) and Sweeney (1992) examine debtors' manipulative behavior before covenant violations. These latter studies find that violations of accounting covenants are costly to debtors, who generally try to manipulate accounting numbers to avoid or defer technical defaults.

The present study also focuses on the event of violation, but from the perspective of creditors. It explains two aspects of creditors' decision process following covenant violations. First, we find that creditors react to actual violations in two distinct ways: they could either waive the violations or could demand certain conditions such as early payment, increase of interest rate, reduction of borrowing base, and so forth.

Second, we also model creditors' decisions either to waive or to call the debt using the option-pricing framework. We hypothesize that the determinants of waiver decisions include the firm's bankruptcy probability and leverage ratio. Moreover, maturity, size, and security of the debt issue involved should also be important factors in the waiver decisions.

The authors acknowledge the helpful comments received from Samir El-Gazzar, Steve Pruitt, An Thakor, Joseph Weintrop, Larry Weiss, and the participants of seminars at Indiana University, Rutgers U versity, and the 1991 annual conferences of the American Accounting Association and the Financial Mana ment Association. We also benefitted from the helpful comments of the anonymous reviewers.

Submitted March 1
Accepted September 1

Empirically, we find that creditors are more likely to grant a waiver to the firm with a lower estimated probability of bankruptcy and a lower leverage ratio. Further, debt issues that are secured or smaller in size are more likely to have violations waived than unsecured or larger issues. The maturity variable, however, is not found a significant determinant of the waiver decisions.

Using the factors identified in this study, managers can assess the probability of receiving a waiver and prepare necessary strategies to ensure the firm's survival. Auditors also can use those factors to assess the possibility of the client's receiving a waiver of covenant violation as part of their evaluation of the firm's ability to continue as a going concern.

Moreover, since debtors prefer waivers to nonwaivers, the prospect of receiving a waiver is likely to influence managerial behavior, including the choice of accounting alternatives. Managers expecting a nonwaiver from creditors would have more incentive to select accounting methods to avoid covenant violations.

Key Words: *Debt covenants, Covenant violation, Technical default, and Waiver.*

Data Availability: *Data used in this study were obtained from public sources. A list of sample firms is available from the first author upon request.*

THE remainder of the paper is organized as follows. In section I, we describe the sample selection process and show the extent to which covenant violations were waived by lenders. In section II, we develop a model to identify the factors that might explain the lender's waiver decision. We report the empirical results in section III and conclusions in section IV.

I. Identification of Firms Violating Accounting Covenants

To identify firms violating accounting-based debt covenants, we searched the "management's discussion and analysis" and "financial footnotes" fields of the Compact Disclosure data base for mentions of "not in compliance," "in violation," "in default of covenants," or "in default of lending, debt, or loan agreements." The 10-K reports were also examined to obtain detailed information about each covenant violation. To assure that the firm's violation had occurred for the first time during the interval covered by this study (1985–1988), we examined the prior year's 10-K report to confirm the nonexistence of violation.

Finally, to concentrate on violations of accounting-based covenants, we excluded firms that were already or simultaneously in default of debt payments or were involved in bankruptcy proceedings when the covenant violation took place. This procedure assures that all 128 firms in the final sample were in compliance with all debt obligations except those involving accounting-based covenants.

For the 128 violators, we checked the "management's discussion and analysis" and footnotes to the financial statements of the violation year to identify whether a firm received a waiver after the violation. Since the events subsequent to annual report date are not considered, the waiver status reported here reflects creditors' initial reactions. Creditors' decisions subsequent to annual report date will be discussed later.

When the waiver was granted with certain conditions, such as acceleration of payments or an increase of interest rates, the violations are classified as not having been waived. In summary, 57 firms are classified as having received waivers and 71 (group III) as having failed to receive a waiver. Of the 57 waivers, 24 (group I) were granted for a limited time period (11 were for less than one year, and 13 were for one year or longer), and 33 (group II) did not have a time limit.

Before we describe the characteristics of the violators, an important *caveat* is in order. Current accounting regulations do not mandate the disclosure of all covenant violations. The SEC's Regulations S-X (Sec. 210.4) requires the disclosure of the facts and amounts of a default if it is not waived or has been waived for a stated period. In addition, *Financial Accounting Standard No. 78* (FASB 1983) requires that debt with covenant violations be classified as a current liability unless a waiver has been granted for more than one year. Thus, management has the option of not reporting violations that were permanently waived. According to DeFond and Jiambalvo (1991), lenders with relatively minor violations who wish to avoid making the disclosure may be able to do so by retroactively renegotiating the terms of the debt agreement. Thus, the sample tends to include relatively serious violations and the actual number of firms receiving a permanent waiver undoubtedly exceeds that reported.

Since this study uses Compact Disclosure as the only data source, our sample cannot be considered as the full population of firms violating accounting covenants in the 1985–1988 period.[1] However, of that data base, most firms that received a limited-time waiver or did not receive a waiver are included, but those receiving a permanent waiver are underrepresented.

Table 1 lists the characteristics of the 128 violators included in the sample, with the waiver versus nonwaiver decomposition for each characteristic. For the number of financial variables violated, more than one half of the firms (68) did not specify which variables were violated. Typically, vague language, such as "certain" or "several financial variables" was used in the footnotes. When specific variables were mentioned, we found 37 violations involved one variable, 17 involved two variables, and 6 involved three or more variables. The variables violated, in order of frequency, were net worth (or tangible net worth), debt to equity, working capital (or current ratio), interest coverage ratio, cash flows, and, in a few cases, profitability.

The breakdown according to the number of creditors associated with each debt issue shows that the vast majority of the debt issues (74) were private loans owed to one creditor; 50 were private loans lent by more than one creditor, and only four debt issues were held by the public.[2] The small number of public debts in the violation sample is consistent with the observation of Smith and Warner (1979) that public debt covenants are less restrictive than private ones perhaps because the cost of renegotiating a public debt agreement is higher.

---

[1] Beneish and Press (1993) also found that the NAARS does not include all firms filing with the SEC, since additional violators were identified from the Dow Jones News Service.

[2] We found that several of the "bonds" or "indentures" are held by only one or a few institutions. These debt issues have the appearance of "public debt," but in essence they are privately held and are classified as "private" in table 1.

Table 1
A Description of the Sample Firms

	Number of Firms for Which Violation Was:	
	Waived	*Not Waived*
Year of First-Time Covenant Violation		
1985	19	19
1986	17	20
1987	13	16
1988	8	16
Exchange Listing		
NYSE	8	14
AMEX	12	20
OTC	37	36
Industry Composition (Based on two-digit SIC codes)		
Industrial equipment (#35)	13	7
Electronic equipment (#36)	5	11
Durable wholesale trade (#50)	4	5
Instruments (#38)	2	6
Miscellaneous manufacturing (#39)	4	2
Dining and drinking (#58)	2	3
Oil and gas products (#13)	0	5
Other industries with less than five firms	27	32
Number of Financial Variables Violated		
One	18	19
Two	7	10
Three	6	0
"Certain" or "several"	26	42
Number of Creditors		
Private debt, one creditor	35	39
Private debt, multiple creditors	20	30
Public debt	2	2

Note: Public debt refers to bonds held by the public. Bonds held by a single or a few institutions are defined as private debt.

II. The Creditors' Waiver Decisions

The Value of Debt and the Waiver Decision

When a debtor is not in compliance with accounting-based covenants, the creditors face a decision either to demand immediate payment or to waive the violation.[3] This decision can be analyzed by using the Black and Scholes (1973) option-pricing framework; that is, the creditors are viewed as owning the firm and selling a call option to the stockholders (see appendix). Without loss of generality, we will base our discussion in the context of a discount bond. At maturity date t^*, the creditors of the discount bond will receive,

---

[3] As will be shown later in table 5, not all the defaults are resolved by immediate payments. They could be resolved by restructuring the debt or by tightening the terms of the loan.

$$\min(V^*,F)=V^*-\max(0,V^*-F), \tag{1}$$

where V^* is the value of the firm at the maturity date, and F is the face amount of the debt. At any time t before maturity, the value of the debt (D) is equal to the value of the firm (V, a random variable) minus the value of the stock, which is a call option with V as the underlying asset. The current value of the debt, D, can be derived from the Black-Scholes option-pricing formula (see the appendix for the mathematical development).

If a waiver is granted when a firm violates accounting-based covenants, the value of the debt will remain as D. If the loan is called, the creditors' expected value of the debt, denoted as D', would become,

$$D'=\min[V',F\bullet\exp(-r\tau)]-OL, \tag{2}$$

where $\tau=t^*-t$, or the time to maturity. Equation (2) indicates that, by calling the loan, the creditors receive the lesser of the firm's value (V') upon the call and the discounted value of the face amount, after subtracting a term representing the opportunity loss of calling the loan (OL). The value of the firm after the call, V', could be lower than V (the value of the firm without the call) since calling the loan will increase the firm's bankruptcy probability. However, the opportunity loss of calling the loan lies primarily in the creditors' loss of a client if the firm is able to fully pay the debt.[4] Therefore, a creditor's decision to waive the violation can be summarized in the following rules:

If $D>\min[V',F\bullet\exp(-r\tau)]-OL$, then waive.

If $D<\min[V',F\bullet\exp(-r\tau)]-OL$, then do not waive.

Since the magnitude of V' is unknown, the rule for a given violation depends on the value of D (relative to the size of the debt, F) and OL. That is, a waiver is more likely to be granted if the relative value of the debt (D/F) and/or the opportunity loss (OL) is higher.

Hypotheses

Several hypotheses can be developed on the basis of the above discussion. First, the probability of bankruptcy should be a factor in the waiver decision. As a firm faces a higher degree of bankruptcy risk, the value of its debt (D) will be lower. In addition, a higher probability of bankruptcy means that the client firm is less valuable to the creditor, which reduces the opportunity cost of losing the client. Thus, a higher degree of bankruptcy risk would make waiver less likely.

H1: A firm with a relatively higher probability of bankruptcy is less likely to receive a waiver from the creditor than one with a lower probability of bankruptcy.

From the determinants of the debt value, several additional hypotheses can be developed. According to the Black-Scholes framework discussed in the appendix, D is a function of the firm's value, its variance, the face amount of the debt, the time to maturity, and the risk-free interest rate. Three of these variables or their combinations are measurable: (1) the face amount of the debt divided by the total value of the firm, which is a typical measure of financial leverage, (2) the time to maturity of the debt, and (3) the face amount of the debt. Based on the mathematical proof in the appendix, the following three hypotheses can be derived:

---

[4] The magnitude of this loss depends on the creditor's ability and the costs to locate a comparable client.

H2: A firm with a higher financial leverage ratio is less likely to receive a waiver from the creditor.
H3: A debt issue with a longer time to maturity is less likely to receive a waiver from the creditor.
H4: The larger the principal amount of a debt issue, the less likely that creditors will waive covenant violations.

The intuitive explanation of hypothesis H3 is that, as the time to maturity increases, the probability of actual default becomes higher.[5] Thus, creditors would have to bear a higher degree of risk by waiving the covenant violation of a longer-term loan. Hypothesis H4 results from the fact that creditors have more to lose in the default of a larger debt issue. Hence, upon covenant violations, the creditors are more likely to adopt measures to protect themselves against these potential losses.

In addition, the security of the debt should have a bearing on the waiver decisions. Stulz and Johnson (1985, 505) show that a secured debt's value is an increasing function of the value of the assets that are pledged as collateral. Thus, debt with collateral should have a higher value of D and makes a waiver more likely than the debt without collateral. This hypothesis can be written as:

H5: A secured debt is more likely than an unsecured debt to receive a waiver from the creditor.

III. Empirical Results

Measurement of Variables

Five variables are necessary to test the five hypotheses. Two are firm-specific (probability of bankruptcy and leverage), and three are debt-specific (maturity, size of debt, and security). All are measured at the end of the fiscal year in which covenant violations took place. For example, if a firm with a December fiscal year-end violated accounting covenants as of December 31, 1985, the variables are measured by using the values as of December 31, 1985.[6]

Previous studies on financial distress provide numerous prediction models that can be used to estimate a firm's probability of failure (see Zavgren 1983 for a review of this literature), but we chose the model reported in Zmijewski (1984, col. 6, panel A of table 3) for its parsimony and freedom from the oversampling bias inherent in many other studies.[7] The probability of bankruptcy ($BKPROB$) is calculated as:

$$BKPROB = N(-4.803 - 3.599NITA + 5.406BLEV - .100CR), \qquad (3)$$

---

[5] A formal proof of this point can be found in Ho and Singer (1982, 397).

[6] We found that most of the covenant violations occurred at the end of the fiscal year. Thus, the creditors' waiver decisions are based, at least partially, on the financial statements for that fiscal year.

[7] We also used a model reported in Ohlson (1980) with virtually identical results.

where *NITA* is net income to total assets, *BLEV* is the leverage ratio based on the book value (total debt divided by total assets), *CR* is the current ratio (current assets to current liabilities), and $N(\bullet)$ is the cumulative probability under the standard normal distribution.[8] The *BLEV* variable is also the leverage variable stated in hypothesis H2.[9] It is the same as that employed in previous studies of the debt covenant hypothesis of accounting choice as a proxy for the tightness of debt covenants (e.g., Holthausen 1981; Leftwich 1981; Lys 1984; among others).

The maturity variable (*MATUR*) is defined as the number of months between the violation date and the maturity date of the debt. If a particular debt issue was granted under a revolving credit agreement, its maturity is defined as zero since the creditors have a periodical (usually annual) option to terminate the credit agreement. If a debt issue is being repaid on an installment basis, its maturity is defined as the time interval between the violation date and the date the next payment is due.[10]

The principal amount (hypothesis H4) is measured as the logarithm of the principal amount of the debt issue for which covenants are violated.[11] It is denoted *DSIZE* and represents the size of the debt issue. Finally, the dummy variable for security (hypothesis H5) is denoted one if the debt is secured, and zero otherwise. Although the market value of the assets serving as security for the debt would provide a finer measure in testing hypothesis H5, such a value is very difficult to determine. Thus, the dichotomous measure of security denoted *SECUR* is used as an indicator for the existence of collateral.

Comparison of Waiver and Nonwaiver Groups

The five variables defined above are compared for three groups in table 2. Group I includes the 24 firms that received temporary waivers, group II includes the 33 firms that received permanent waivers, and group III includes the 71 firms that did not receive a waiver. The comparison between groups I and II (from both *t*- and Wilcoxon *Z*-tests) shows that the two groups do not have statistically different mean values for the five variables, suggesting that the firms receiving permanent waivers have the same profile as those receiving temporary waivers. Accordingly, they are combined for the purpose of the next comparison.

The comparison between group III versus groups I and II combined shows that firms receiving waivers have significantly lower mean values for *BKPROB*, *BLEV*, and *DSIZE*, consistent with hypotheses H1, H2, and H4. Overall, table 2 shows that all three groups have rather high mean values for *BLEV* relative to typical firms without violation of debt covenants. This difference suggests that firms violate debt covenants when

---

[8] According to SFAS No. 78, a long-term debt should be classified as current if the covenant violation is not waived for more than a year. Following previous bankruptcy studies, the *CR* variable in equation (3) is based on the data reported in financial statements, and thus reflects the reclassification from long-term to current. In the alternative tests not reported here, we classified the debts back from current to long-term if they had been reclassified. We found no change in the results.

[9] We also obtained the leverage ratio by using the market value of the equity (the book value of the debt divided by the market value of the equity plus the book value of the debt) instead of the book value of the debt. This measure of leverage had a similar, but weaker, association with the waiver decision in the logit regressions.

[10] In a few instances, the installment payment interval was not specified in the footnotes. The *MATUR* variable for these cases is defined as zero.

[11] If the covenants include a cross-default clause, all the debt issues involved are included in the *DSIZE* variable.

Table 2
Financial Profiles of Sample Firms

Waiver Status	*Variables*[a]				
	BKPROB	BLEV	MATUR	DSIZE	SECUR
Group I: Time-limited waiver (N=24)					
Mean	0.254	0.684	7.00	1.968	0.750
Median	0.143	0.686	1.00	1.594	
Std. Dev.	0.251	0.122	11.52	1.717	
Group II: Permanent waiver (N=33)					
Mean[b]	0.274	0.668	15.06	1.776	0.788
Median	0.195	0.704	6.00	2.054	
Std. Dev.	0.279	0.224	23.45	1.282	
Test of Differences Between Groups I and II					
t-statistic[c]	−0.291	0.352	−1.550	0.484	0.263
Wilcoxon Z-test	−0.137	0.331	−1.333	0.073	
Group III: Nonwaiver Group (N=71)					
Mean[b]	0.429	0.742	12.37	2.424	0.620
Median	0.369	0.738	3.00	2.197	
Std. Dev.	0.332	0.242	20.49	1.553	
Test of Differences of Groups I and II combined vs. Group III					
t-statistic[c]	−3.094**	−1.769*	−0.196	−2.103*	1.900*
Wilcoxon Z-test	−2.999**	−1.628*	0.449	−2.122*	

[a] Definition of Variables: *BKPROB* is the probability of bankruptcy calculated from the model in Zmijewski (1984); *BLEV* is book value of debt divided by book value of total assets; *MATUR* is the debt issue's time to maturity (in months); *SECUR* is equal to 1 if the debt is secured, 0 otherwise; *DSIZE* is the logarithm of the total amount of debt in technical default.

[b] Since *SECUR* is a dichotomous variable, its mean is the proportion of firms with an observation of 1 in that variable and the entries for *t*-statistics are the Z-statistics testing whether the proportions are the same for the two groups.

[c] *t*-statistic tests, for each variable, whether the mean of the waiver group is larger than that of the nonwaiver group.

* Significant at the 0.05 level (one-tailed test).
** Significant at the 0.01 level (one-tailed test).

they are in financial difficulty. However, the firms that were granted waivers had relatively better financial positions than their counterparts in the nonwaiver group.

The debts in the waiver group have shorter average maturity, but it is not significantly different from that of the nonwaiver group. In addition, although most debt issues were secured, the waiver group has a significantly higher percentage of secured debt (at the 0.01 level) than the nonwaiver group, which is consistent with hypothesis H5.

Table 3 shows that *BKPROB* is highly correlated with *BLEV* (with a coefficient of 0.77). This is not surprising since *BLEV* is involved in the calculations of *BKPROB*. For this reason, the logit models presented below employ these two variables separately. All

Table 3
Correlations of Determinant Variables

	BKPROB	BLEV	MATUR	DSIZE
BLEV	0.77**			
MATUR	0.12	0.19*		
DSIZE	0.14	0.34**	0.15*	
SECUR	0.09	0.12	−0.14	0.03

Note: The variables are defined under table 2.

* Significant at the 0.05 level (one-tailed test).
** Significant at the 0.01 level (one-tailed test).

other bivariate correlation coefficients are not particularly high, which allows a less noisy interpretation of the remaining variables.

Results of Logit Regressions

The five hypotheses can be tested by the following logit models:

$$\text{Model I: } \log\left[\frac{P(waiver)}{1-P(waiver)}\right]=\beta_0+\beta_1 BKPROB+\beta_3 MATUR+\beta_4 DSIZE+\beta_5 SECUR,$$

and

$$\text{Model II: } \log\left[\frac{P(waiver)}{1-P(waiver)}\right]=\beta_0+\beta_2 BLEV+\beta_3 MATUR+\beta_4 DSIZE+\beta_5 SECUR,$$

where $P(waiver)$ is the probability that the creditor will waive a covenant violation, and all other variables are as previously defined. Model I employs *BKPROB*, while Model II uses *BLEV*. According to hypotheses H1 through H5, the signs of the coefficients are expected to be: $\beta_1<0$, $\beta_2<0$, $B_3<0$, $\beta_4<0$, and $\beta_5>0$.

The logit models were estimated with a binomial classification of the dependent variable; that is, the violated debt covenants of the sample firms are classified as either waived or nonwaived. No attempt was made to further classify the nonwaiver group according to severity of consequences of failing to obtain waivers because the independent variables are measured as of the end of the fiscal year during which the violation took place. In many cases, the eventual resolutions of violations were disclosed in the annual reports for the year subsequent to the violations, because the negotiation is a time-consuming process. To further decompose the nonwaiver group would require the use of information that was made known subsequent to the measurement time of the independent variables.

The logit regression results are shown in table 4, together with the hypotheses corresponding to each independent variable and the expected sign of the coefficients. The overall chi-square statistics for both models are significant at conventional levels.

The results from Model I suggest that the waiver decision is positively related to the security of the debt, and negatively related to the bankruptcy probability and the value

Table 4
Logit Regressions Explaining the Waiver Decisions
(*t*-statistics in parentheses)

Variable	*Model I (BKPROB)*	*Model II (BLEV)*	*Model III (Excluding Permanent Waivers)*	*Model IV (Using Later Information)*	*Expected Sign*
Intercept	0.186	0.477	−0.563	−0.97	
	(0.40)	(0.71)	(−0.95)	(−1.71)	−
BKPROB	−1.984		−2.078	−1.547	−
(H1)	(−2.89)**		(−2.14)*	(−1.90)*	
BLEV		−1.298			−
(H2)		(−1.38)*			
MATUR	0.005	0.004	−0.008	0.013	−
(H3)	(0.45)	(0.57)	(−0.53)	(1.66)*	
DSIZE	−0.241	−0.218	−0.127	−0.363	−
(H4)	(−1.82)*	(−1.62)	(−0.77)	(−2.23)*	
SECUR	1.027	0.907	0.771	1.231	+
(H5)	(2.33)**	(2.13)*	(1.33)	(2.18)*	
Model Chi-square	17.86**	10.41*	9.44*	14.35**	
Pseudo-R	0.237	0.117	0.116	0.214	
Number of waivers	57	57	24	31	
Number of nonwaivers	71	71	71	92	

Note: The dependent variable is defined as 1 if a waiver is received, and 0 otherwise. The independent variables are defined under table 2. Model III is the same as Model I except that the firms receiving a permanent waiver are excluded. Model IV uses the information available one year after the violation to define the waiver status. The 31 firms whose debt terms were not changed in the year after the covenant violation are defined as "waived"; the remaining 92 firms are defined as "nonwaived" (see table 5).

* Significant at the 0.05 level (one-tailed test).
** Significant at the 0.01 level (one-tailed test).

of the debt. On the other hand, the maturity variable is not found to be a significant determinant of waiver decisions.[12]

Using leverage in Model II produces qualitatively the same results as those for Model I using bankruptcy probability. The significance level for the coefficient on *BLEV*, however, is much lower than that of *BKPROB*. This is not surprising since *BKPROB* incorporates more dimensions of the firm's financial condition than the *BLEV* variable.

Table 4 presents the results for Model III, which is the same as Model I but excludes the 33 firms receiving a permanent waiver. As noted in the discussion on sample selection, current disclosure requirements lead to an underrepresentation of firms receiving a permanent waiver. Although the signs of all the coefficients except *MATUR* are the same as in Model I, Model III's overall logit chi-square, pseudo-R and the significance level of each coefficient are all below those in Model I.

---

[12] The lack of significance could be attributable to a low level of inherent variation. As shown in table 2, the median values of *MATUR* are very small for each group, indicating that most of the debts have very short maturity.

Later Development of Waivers and Nonwaivers

In the above analysis, waiver status was identified by using the information available in the annual reports for the year of violation and creditors' actions after the annual report date were not considered. In this respect, waiver status in the prior analysis represents the creditors' initial reactions to violations.

It is possible that creditors might grant waivers initially, then demand certain additional conditions due to either the original or a subsequent violation. To identify the later developments of covenant default, we examined the footnotes of the year following the violation. By comparing the terms (such as borrowing bases, interest rates, security, maturity, etc.) of the relevant debt issues reported in the violation year and the following year, renegotiation of terms can be identified. In table 5, the later developments of covenant violations are reported for the two waiver groups (temporary and permanent) and the nonwaiver group.[13]

For the permanent waiver group, creditors in four cases (one firm failed to pay and filed for bankruptcy; the other three were able to pay) demanded payments the following year. One firm's debt was restructured, one firm's debt was converted to a demand loan, and two firms' debt terms were tightened (including a reduction of borrowing base, an increase in interest rate, additions to collateral, etc.). In 21 cases, we found the terms of the debt unchanged during the year following the violation.[14]

For the temporary waiver group, payment was demanded by creditors in four cases (two firms paid and two were not able to pay). In addition, two firms' debt issues were converted into equity, two firms' debt was restructured, and five firms' debt terms were tightened. In ten cases, the terms of the debt were found unchanged during the year following the violation.

As to the nonwaiver group, creditors demanded payment in 31 cases (21 firms paid, and ten were unable to pay). Eight firms' debt issues were converted to equity or paid back by issuance of equity stocks. Seven firms' debt issues were restructured (two of these are classified as a "troubled-debt restructuring," in which the creditors settled for less than the principal amount). Six firms' debt issues became demand loans. Finally, the debt terms were tightened in 19 cases.[15]

In summary, three observations can be made. First, not all firms that initially received waivers were completely out of trouble in the year following the violation. Second, of the 31 firms whose debt terms were found to be unchanged one year following the debt covenant violation, it appears that the permanent waiver group has a much higher percentage of firms without known costly consequences (21 out of 29 vs. ten out

---

[13] For the nonwaiver group, some "later developments" were already stated in the annual reports for the violation year. In those cases, the creditors refused to waive the violation, but did not take any actions immediately. We define such cases as nonwaiver.

[14] In several cases, we found that the financial ratios required in the debt agreements were relaxed. This change is not defined as "terms tightened" in table 5 since it does not impose any costs on the part of the debtor.

[15] The credit terms tightened include changes in borrowing base, interest rates, and collateral. In three cases, firms granted the creditors stock warrants as a condition of a waiver.

Table 5
Later Developments of Covenant Violations

	Firms With Permanent Waiver	*Firms With Temporary Waiver*	*Firms Without Waiver*
Later Developments:			
Payment demanded by creditors:			
Failed to pay	1	2	10
Paid in full or in part	3	2	21
Debt converted into equity, or issued stock to pay debt	0	2	8
Debt restructured	1	2	72
Debt becomes demand loan	1	0	6
Terms of debt tightened	2	5	19
No changes in the terms of debt	21	10	0
Later development unknown (annual report unavailable for following year)	4	1	0
Total	33	24	71
Source of Later Development Information (Excluding the firms whose debt terms were not changed the year after covenant violation):			
Annual reports of the covenant violation year	0	0	50
Annual reports of the year after violation	8	13	21

Note: Restructuring includes one or more of the following: rescheduling of payment dates, changes of interest rates, provision of collateral, etc. All restructurings were unfavorable to the firms except for two "troubled-debt restructurings," in which the creditors settled the debt for less than the loan amount. The terms tightened include provision of collateral, reduction of borrowing base, increase of interest rate, and granting of stock warrants. Relaxation of accounting covenants is not considered change in the terms of debt.

of 23 firms given temporary waivers). Finally, it is apparent that the consequences of violating accounting covenants are much more severe for the nonwaiver group than for the two waiver groups.

For the permanent waiver group, it is possible that the subsequent nonwaivers could be due to repeated violation in the following year. For the temporary waiver group, creditors could demand payment or additional conditions after the waiver expires. Unfortunately, firms typically do not reveal whether the nonwaiver is due to original or subsequent violation. As a result, the waiver decisions reported in table 5 could be in response to a second violation, which could differ from the decisions made upon the first violation. The first three logit models reported in table 4 are based on the creditors' decisions with respect to the first violation. As a comparison, Model IV considers the later consequences of violations. Specifically, the 31 firms whose debt terms were not altered are defined as "waived," and the remaining 92 firms with costly consequences are defined as "nonwaived."[16] The results are very similar to those of Model I, except that *MATUR* is significantly positive in Model IV, but insignificant in Model I.[17]

All the information on later developments is identified from the annual reports one year following the covenant violation. Thus, Model IV's independent variables are measured at the end of the violation year, but its dependent variable (waiver status) is based on subsequent information. Despite this inconsistency, Model IV shows that the variables identified in this study are still able to explain both creditors' initial and later reactions to covenant violations.

IV. Conclusions

The article examines creditors' reactions to violations of accounting debt covenants by identifying the factors underlying creditors' waiver decisions. We found that, among violations reported by the firms, a waiver is more likely to be granted to a firm with a lower estimated probability of bankruptcy and a lower leverage ratio. In addition, secured or small debt issues have a better chance of receiving violation waivers. Similar results hold when the information one year after the violation was used to define waiver status.

The above results are subject to an important *caveat*. Namely, firms receiving permanent waivers are expected to be underrepresented in our sample since many of these waivers could go unreported according to current disclosure requirements concernng debt covenant violations. How this sampling problem affects our results can only be conjectured. According to DeFond and Jiambalvo (1991), the unreported violations might be relatively minor. Thus, inclusion of the unreported firms in the waiver group could make that group's financial profile even better, leading to a larger difference between the waiver and the nonwaiver groups. As a result, the omission of unreported violations seems to have a downward bias against finding difference between the two groups.

The findings of this study can be useful to at least three major user groups of accounting information. First, commercial lending officers, who often must make waiver decisions, can use the systematic factors identified in this study as a decision aid. Second, managers of firms that are about to be in violation of an accounting covenant could use these findings to assess the probability of receiving a waiver and to prepare necessary strategies that ensure the firm's financial survival. Finally, auditors are required by SAS No. 59 (AICPA 1988) to "evaluate whether there is substantial doubt about the entity's ability to continue as a going concern . . ." (par. 2). Several factors that might be considered in making this evaluation are suggested, including loan default and debt restructuring (par. 6). If a client is involved in a covenant violation, the factors of waiver decisions reported in this study should be helpful for the auditors' evaluation of going-concern uncertainties.

The findings in this study also provide additional insight into the "debt covenant hypothesis" of accounting choices, which postulates that management alters accounting methods to move the firm away from debt covenant restrictions (see Watts and Zimmerman 1990 for a survey).[18] Since debtors prefer waivers to nonwaivers, the prospect of a waiver is likely to influence managerial behavior, including the choice of accounting alternatives. Previous studies have shown that managers have an incentive to alter accounting methods when the firm is close to violating accounting covenants. From this study, it can be inferred that such an incentive is reinforced when a waiver is not expected from the creditors, and vice versa. Thus, the factors in the waiver decisions should be helpful for future research that attempts to further explain managers' accounting choices.

---

[16] The total number of firms used in Model IV is 123, five less than in Models I through III, because of the unavailability of 10-Ks in the year after violation.

[17] According to hypothesis H3, the maturity variable is expected to have a negative sign, since longer-term debts are less likely to receive a waiver according to our theoretical developments. The unexpected sign could be due to the imprecise measurement of the *MATUR* variable. In particular, since a revolving loan's maturity is unknown, it has been defined as zero.

Several questions not answered in this study could also prove fruitful topics for future research. First, there were very different consequences for the nonwaiver group. Some (such as payments demanded) are more severe than others (such as tightening terms). Beneish and Press (1993) attempt to measure the monetary costs of those consequences. In our study, we treated all the nonwaiver firms as one group; further studies could attempt to explain the different consequences more carefully. Second, the study employs only the variables available in financial reports. Inclusion of additional variables to which creditors have access could increase the explanatory power of the models. Finally, the factors identified in this study are largely derived from the option-pricing framework. Alternative methodologies might generate additional factors and better understanding of creditors' waiver decisions.

Appendix

The Black-Scholes Model as a Basis for Hypotheses H2–H4

According to the Black-Scholes option-pricing formula, the value of the debt described in equation (1) can be written as:

$$D=V\cdot N\left[\frac{-\ell n(V/F)-(r+\sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right]+F\cdot \exp(-r\tau)\cdot N\left[\frac{\ell n(V/F)+(r-\sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right] \tag{A1}$$

where $N(\cdot)$ is the cumulative standard normal distribution function, V is the firm's value, F is the face amount of debt, σ^2 is the instantaneous variance of V, r is risk-free interest rate, and τ is the time to maturity.

Hypothesis H2, stated in section II, is based on the partial derivative of D/F (the value of debt relative to the face amount of debt) with respect to F/V (the debt-to-assets ratio):

$$\frac{\partial(D/F)}{\partial(F/V)}=-(F/V)^2\left[\frac{\partial(D/F)}{\partial(V/F)}\right]=-(F/V)^2\cdot N\left[\frac{-\ell n(V/F)-(r+\sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right]<0. \tag{A2}$$

From equation (A2), it is clear that the leverage ratio affects the value of D/F adversely, and hence hypothesis H2. In addition, the partial derivative of D/F with respect to τ can be shown to be negative. This leads to hypothesis H3. Finally, the following expression shows that F and D/F are negatively related, hence providing the theoretical basis for hypothesis H4:

$$\frac{\partial(D/F)}{\partial F}=-(V/F^2)\cdot N\left[\frac{-\ell n(V/F)-(r+\sigma^2/2)\tau}{\sigma\sqrt{\tau}}\right]<0. \tag{A3}$$

---

[18] Since debt contract information is not easily observable, most of these studies employ financial leverage ratios (e.g., debt/equity ratio, see Holthausen 1981 and others) as a proxy for the existence or tightness of the covenant restrictions. For tests of the validity of this approach, see Duke and Hunt (1990) and Press and Weintrop (1990), among others. In addition, Begley (1990) provides review on this line of research.

References

American Institute of Certified Public Accountants. 1988. *Statement of Auditing Standards No. 59: The Auditor's Consideration of an Entity's Ability to Continue in Existence*. New York: AICPA.

Begley, J. 1990. Debt covenants and accounting choice. *Journal of Accounting & Economics* 12 (January): 125–39.

Beneish, M.D., and E. Press. 1993. Costs of technical violation of accounting-based debt covenants. *The Accounting Review* 68 (April): 233–57.

Black, F., and M. Scholes. 1973. The valuation of options and corporate liabilities. *Journal of Political Economy* 81 (May–June): 637–59.

DeFond, M., and J. Jiambalvo. 1991. Debt covenant violation and manipulation of accruals. Working paper, University of Southern California, Los Angeles, CA.

Duke, J. C., and H. G. Hunt III. 1990. An empirical examination of debt covenant restrictions and accounting-related debt proxies. *Journal of Accounting & Economics* 12 (January): 45–64.

Financial Accounting Standards Board. 1983. *Statement of Financial Accounting Standards No. 78: Classification of Obligations That Are Callable by the Creditor*. Stamford, CT: FASB.

Ho, T., and R. F. Singer. 1982. Bond indenture provision and the risk of corporate debt. *Journal of Financial Economics* 10 (December): 375–406.

Holthausen, R. W. 1981. Evidence on the effect of bond covenants and management compensation contracts on the choice of accounting techniques: The case of the depreciation switch-back. *Journal of Accounting & Economics* 3 (January): 73–109.

Leftwich, R. W. 1981. Evidence on the impact of mandatory changes in accounting principles on corporate loan agreements. *Journal of Accounting & Economics* 3 (January): 3–36.

Lys, T. 1984. Mandated accounting changes and debt covenants: The case of oil and gas accounting. *Journal of Accounting & Economics* 6 (January): 39–65.

Ohlson, J. A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18 (Spring): 109–31.

Press, E. G., and J. B. Weintrop. 1990. Accounting-based constraints in public and private debt agreements: Their association with leverage and impact on accounting choice. *Journal of Accounting & Economics* 12 (January): 65–96.

Securities and Exchange Commission. 1988. *SEC Accounting Rules*. Chicago: Commerce Clearing House.

Smith, C. W., Jr., and J. B. Warner. 1979. On financial contracting: An analysis of bond covenants. *Journal of Financial Economics* 7 (June): 117–62.

Stulz, R. M., and H. Johnson. 1985. An analysis of secured debt. *Journal of Financial Economics* 14 (December): 501–21.

Sweeney, A. P. 1992. Debt-covenant violations and manager's accounting-choice strategies. Working paper, Harvard University, Boston, MA.

Watts, R. L., and J. L. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice-Hall.

——, and ——. 1990. Positive accounting theory: A ten year perspective. *The Accounting Review* 65 (January): 131–56.

Zavgren, C. 1983. The prediction of corporate failure: The state of the art. *Journal of Accounting Literature* 2 (Spring): 1–37.

Zmijewski, M. E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22 (Supplement): 59–86.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 233–257

Costs of Technical Violation of Accounting-Based Debt Covenants

Messod D. Beneish
Duke University
Eric Press
Temple University

SYNOPSIS AND INTRODUCTION: Costs associated with the violation of accounting-based covenants in debt agreements are presumed to be material by both accounting regulators and researchers. The Financial Accounting Standards Board, for example, delayed the implementation of its pronouncement on pension reporting, SFAS No. 87, for two years to allow firms sufficient time to "renegotiate or to obtain waivers of provisions of some legal contracts" (FASB 1985, par. 260). Numerous studies in accounting research hypothesize that it is costly for firms to violate accounting covenants in debt agreements, and this supposition figures in research on such issues as the economic impact of mandated and voluntary accounting changes (see, e.g., Holthausen 1981; Leftwich 1981; Lys 1984) and the determinants of accounting choice (see, e.g., Trombley 1989; Zmijewski and Hagerman 1981). Although research in financial economics has studied some of the costs shareholders bear when there are debt service defaults or bankruptcy filings, the costs associated with technical violation—the violation of covenants other than debt service—have not been documented.

This study investigates the costs of technical violation for a sample of 91 firms that violated accounting-based covenants in debt agreements between 1983 and 1987. The sample includes firms for which the technical violation was sufficiently material to merit disclosure. We provide direct evidence of refinancing and restructuring costs by examining changes in terms of debt agreements, and changes in investing and financing

We thank Celal Aksu, Andy Alford, Larry Brown, Jennifer Francis, Stuart Gilson, Bob Hagerman, Bob Holthausen, Tom Lechner, Vic Pastena, Mike Rozeff, Kevin Sachs, Amy Sweeney, Joe Weintrop, Jerold Zimmerman, two anonymous referees, and participants in accounting workshops at Duke, Georgetown, Rutgers, SUNY-Buffalo, and Temple for comments on this manuscript. We are especially indebted to Vic Bernard for advice and encouragement. Pat Williamson of Disclosure, Inc. dispensed invaluable help in obtaining documents; and Elizabeth Corona, Meggan Kinum, John Tortora, and Lisa Williams provided competent research assistance. Much of the work for this article was completed while the second author was visiting at Duke.

Submitted February 1991.
Accepted August 1992.

decisions. Refinancing costs arise because lenders raise interest rates on loans and notes following violation. We estimate that increased interest costs resulting from violation range between 0.84 and 1.63 percent of the market value of sample firms' equity. Restructuring costs stem from lenders' demands for partial or full repayment. Nearly half the sample firms either refinanced their debt or divested assets within one year of violation, stating that the proceeds were to reduce the outstanding balances of violated debt agreements. We estimate that the costs of restructuring debt represent an average of 0.37 percent of sample firms' market value of equity. We also present some evidence that there are costs associated with modifying operations, although we cannot estimate their magnitude; lenders' repayment demands impose restructuring costs by forcing firms to eliminate profitable investment projects.

In addition to these costs, increased lender control is an important effect of technical violation. We observe that lenders add numerous new covenants. Interestingly, few of these are accounting-based, which suggests that only slight adjustments to accounting-based monitoring are required. The majority of new covenants consists of restrictions on investing and financing to prevent further dissipation of assets.

We consider whether the costs of technical violation vary according to lender response. We find that the costs are lower for firms that can obtain a waiver than for those that cannot. More important, the evidence suggests that lenders often extract fees and concessions from violators in exchange for granting waivers.

This is one of the first studies to substantiate that technical violation of accounting-based covenants is costly. Depending on the assumptions made, the average costs we estimate range between 1.2 and 2 percent of market value of equity; alternatively, the losses represent between 4.4 and 7.3 percent of the outstanding balances of the violated debt agreements. Evidence on the costs of technical violation is relevant to researchers who attribute economic consequences to changes in debt covenant slack and the likelihood of violating accounting-based covenants. Furthermore, by showing that leverage proxies for the magnitude of some of the costs imposed by technical violation, we justify the use of this surrogate in accounting research.

Key Words: *Debt covenant violation, Technical default, Renegotiation, Waiver, Private lending agreements.*

Data Availability: *The sample of violating firms is available from either author.*

IN section I, we discuss costs of technical violation and relate the study to the literature on workouts and bankruptcy. We discuss the procedures used to identify sample firms and collect data in section II. We describe our method in section III and present our findings in section IV. Our conclusions appear in section V.

I. Costs Imposed by Technical Violation

Debt agreements require borrowers to inform lenders of any breach of the terms in their lending agreements.[1] A default on debt service or breach of a covenant gives rise to a potential "event of default," as defined in *Commentaries on Indentures* (ABF 1971, sec. 5-1). Failure to cure the breach within a grace period (usually 30 days) results in an event of default, which gives lenders the option to seek remedies such as modifying the terms of lending agreements or terminating the loan credit. Renegotiation between lender and borrower typically follows a technical violation.[2]

A primary consequence of technical violation is that it gives lenders the option to exercise contractual rights that potentially impose renegotiation, refinancing, and restructuring costs.[3] Renegotiation costs are the costs of rewriting contracts and typically include expenditures for arranging a workout of loan defaults (e.g., managers' time and fees for attorneys and accountants).[4] These costs have been documented in several studies of bankruptcy. Prior research commonly identifies renegotiation costs as the direct cost of financial distress; estimates of these costs for bankrupt firms range between 3 and 7 percent of the value of the firm (Ang et al. 1982; Warner 1977; Weiss 1990). In our sample, which mainly contains technical violations of private lending agreements (87 of 91 firms), costs of renegotiation are likely to be neither great (Leftwich 1981) nor well-documented (no court records are required because all negotiation is private), and we focus instead on refinancing and restructuring costs.

Refinancing costs are incremental costs of borrowing, and restructuring costs are associated with changing capital structure and modifying operations (e.g., costs of exchanging debt and the decline in performance associated with forced divestiture of assets). Prior research on debt service default and bankruptcy usually designates these costs as "indirect" and treats them as unmeasurable because of data limitations (Gilson 1990; Weiss 1990), but the approach in this study allows us to estimate these costs by comparing debt contract terms pre- and post-violation. Further, since firms in technical violation need not undergo structural changes as drastic as those in reorganization proceedings, we also furnish evidence on how lender demands affect managers' investing and financing decisions post-violation.

---

[1] Absent a covenant violation, firms provide a certificate of compliance along with quarterly and annual financial statements. The certificate is signed by the principal financial officer, who states that nothing has come to his or her attention regarding any event of default or breach of covenant.

[2] *Commentaries on Indentures* (ABF 1971, 2) notes there is "no body of statutory or common law that protects the holder of unsecured debt securities," so the rights of a lender are largely set within the context of the contract negotiated with the borrower. There is no set response to private agreement violation, and procedures for notification of violation vary across contracts and lending institutions. In the words of a senior bank executive, "Although there are some . . . standard covenants on debt securities, other than an 'Appalachian Oral History' approach among negotiators and their lawyers, limited documentation of covenant standards or critical patterns of default seems to exist" (private correspondence, September 1991).

[3] We do not consider implicit contracting costs that could be imposed by customers, vendors, or employees, e.g., higher costs if customers and distributors decline to reorder products or if employees seek more secure positions and resign more frequently.

[4] Other authors, such as Leftwich (1981), include refinancing and restructuring costs with renegotiation costs.

II. Data

Sample Selection

Generally accepted accounting princples (GAAP) do not directly require disclosure of covenant violations, nor is there a systematic public record of technical violations. However, SFAS No. 78 (FASB 1983) and Emerging Issues Task Force Release 86-30 (FASB 1986) require disclosure of the circumstances of violation when long-term debt is reclassified as a current liability. SAS No. 59 (AICPA 1989) indicates that default or lack of compliance with covenants requires a going-concern warning and violation disclosure. Regulation S-X of the Securites and Exchange Commission (SEC 1988, sec. 210.4-08) prescribes that "any breach of covenant of *a*[*n*]... indenture or agreement, which... exist[s] at the date of the most recent balance sheet being filed and which has not been subsequently cured, shall be stated in the notes to the financial statements." As a consequence of these rules, financial statements reflect material and uncured violations of debt covenants.

We searched annual reports using Compact Disclosure and NAARS to identify debt covenant violators. Using key words such as "violation," "waiver," "covenant," "compliance," and "default," we identified 202 cases of potential technical violation during fiscal years ending between 1983 and 1987.[5] The sample was obtained by using the following criteria, which are summarized in table 1.

1. The annual report must disclose that during the fiscal year of initial violation, the firm is in technical violation. We eliminated 25 firms that were either not in compliance with environmental regulations or were financial institutions not in compliance with reserve requirements.
2. The contract breach must only be technical violations. We eliminated 34 firms in default of interest or principal payments or filing under Chapter 11. Two of these firms were jointly in technical violation and default of debt service; a third firm was jointly in technical violation and filing under Chapter 11.
3. The first episode of violation must occur during our sample period, conditional upon the firm's not having been in violation in the prior year. We read all firms' financial statements for the year prior to violation to ascertain that the firm was not in violation then. Thus, our firms are first-time violators within the 1983–1987 period. We eliminated 22 firms with violations outside this period.
4. A firm is included in our sample only once, and we used the earliest episode of technical violation. We eliminated 21 instances of repeat technical violation (20 of which occurred jointly with debt service default or Chapter 11 filings).
5. A firm's Form 10-K and annual report must be available in university collections. This was necessary for two reasons. First, we confirmed each violation prospect against a microfiche copy of the firm's Form 10-K. Second, identifying publicly available agreements required us to read item 14 of each firm's Form 10-K. The university collections available to us did not contain NASDAQ firms, and we eliminated 17 OTC firms.

We also searched the Dow Jones News Service (DJNS) on our key words. This complements the search on Compact Disclosure and NAARS since the coverage of finan-

---

[5] We used NAARS to identify firms in 1983 and 1984 since Compact Disclosure begins in 1985.

Table 1
Selection Criteria for a Sample of Firms in First-Time Technical Violation Between 1983 and 1987

Compact Disclosure and NAARS search:	
Cases of potential violation	202
Minus:	
False positives	
Firm not in compliance with environment regulations	15
Firm a financial institution (SIC 6000–6350) not in compliance with reserve requirements	10
Firm defaults on debt service or files Chapter XI	34
First violation is outside the 1983–1987 period	22
Firm is a repeat violator, 1983–1987	21
OTC firms for which Forms 10-K and annual reports are not readily available[a]	17
Remaining firms	83
Dow Jones News Service search:[b]	
Cases of potential violation not identified by Compact Disclosure or NAARS	13
Minus:	
OTC firms for which Forms 10-K and annual reports are not readily available	5
Remaining firms sampled from Dow Jones News Service	8
Total firms sampled	91

[a] Our searches require that Forms 10-K be available. The university collections readily accessible to us contained neither OTC firms' Forms 10-K nor their annual reports.

[b] We searched the Dow Jones News Service because the financial coverage in both Compact Disclosure and NAARS is incomplete, as explained in footnote 6.

cial statements from these sources, while extensive, is not complete.[6] The DJNS search identified 13 additional cases of technical violation, of which we retained eight firms having Forms 10-K available. Our final sample of 91 firms had violations in debt agreements that totaled $3.4 billion. The mean outstanding balance in the violated agreements was $37.6 million, with a median of $20 million. The average size of these debt agreements equals 27.3 percent of market value of equity and 13.6 percent of total debt, when measured at the end of the fiscal year of violation.

The Covenants That Are Violated

We determined that 74 of the 91 sample firms had violated various accounting-based covenants by examining the financial statements and the affirmative and nega-

---

[6] Financial coverage (annual reports, Forms 10-K) of firms in both Compact Disclosure and NAARS is incomplete. Compact Disclosure's annual report footnotes are only available for 75 to 90 percent of New York and American Stock Exchange firms in the period 1985–1987, and NAARS does not ensure that all its requests for annual reports from traded firms are followed up (Press 1988).

Table 2
Distribution of Technical Violation of Accounting-Based Covenants

Covenant	*Number of Firms With Violation*[a]	*Percentage of All Violations*[b]
Tangible net worth	54	42.8
Working capital or current ratio	37	29.4
Leverage	24	19.0
Minimum earnings level or prohibition against losses	5	4.0
Interest coverage	4	3.2
Cash flow from operation	1	0.8
Cash flow to total debt	1	0.8
Paid dividends in excess of restrictions on earnings	0	0.0
Total accounting-based violation	126	100.00

[a] It was possible to identify the covenants violated from the debt agreements and financial statements of 74 firms. For the remaining 17 firms in our sample, we could not identify the violations, but ensure that debt service default is excluded by eliminating firms with missed interest on principal payments (cf. table 1).

[b] Of the 130 violations observed, 126 are accounting-based. The four nonaccounting-based violations (two instances each of excess capital expenditures and borrowing beyond base limits) occur jointly with accounting-based violations.

tive covenants of violated debt agreements.[7] For example, Publicker's 1984 Form 10-K discloses that the "Company has notified its term loan . . . lenders that a technical event of default has occurred. The default related to covenants involving the Company's minimum net worth and the ratio of indebtedness to tangible net worth." For the remaining 17 firms, financial statements do not identify the violated covenants, nor are there publicly available debt agreements to which we can refer. However, we have reasonable assurance that accounting-based covenants are violated. Semtech's 1985 Form 10-K discloses that "certain financial covenants . . . had not been met at February 2, 1985"; and Parker Drilling's 1984 Form 10-K states that its wholly owned subsidiary "was not in compliance with certain financial ratios required under a credit agreement with a group of banks." Thus, it is likely that the violated financial covenants that we cannot observe for the 17 firms are accounting-based, given that we eliminated firms in debt service default from the sample.

The observed incidence of technical violation includes one firm in violation of four covenants, 13 firms in violation of three covenants, 27 firms in violation of two covenants, and 33 firms with one violation. This represents a total of 130 covenant violations, of which 126 are accounting-based. The four nonaccounting-based violations (excess capital expenditures and borrowing in excess of allowable base) occurred jointly with accounting-based violations. Thus, the costs we examine in this study originate mainly in violations of accounting-based covenants, which permits us to assess whether negative economic consequences arise from violation of accounting-based covenants, as hypothesized, but not documented, in prior research.

As table 2 shows, of the 126 accounting-based violations we can identify using financial statements and debt agreements (across 74 firms), tangible net worth, working capital or current ratio, and leverage (restrictions on debt levels relative to assets or shareholders' equity) are the most frequently violated accounting covenants. There are also violations on earnings requirements, interest coverage, and cash flow requirements (11 incidents). We found no violations for payment of dividends in excess of restrictions.

---

[7] Affirmative covenants contain pledges by a borrower to take specific actions, and agreements to keep financial ratios and levels within specified limits. Negative covenants proscribe activities and also specify ratio limits (Press and Weintrop 1990).

Causes of Violation

Adverse changes in financial statement numbers lead to technical violations of accounting-based covenants. Thus, financial distress and income-decreasing accounting changes can induce violation, and we investigate both possible causes.

To assess the part played by financial performance in inducing technical violation, we contrast characteristics of technical violators with those of nonviolators. In table 3, we compare our sample firms' financial characteristics to those of industry-matched nonviolating firms with data available on the COMPUSTAT primary, supplementary, tertiary, and full coverage files. Nonviolators enter the comparison sample as members of the 34 two-digit industries to which violators belong.[8] The comparison between violators and industry-matched nonviolators is made one year prior to technical violation to reduce distortions caused by violation.

As shown in table 3, violators present a profile consistent with financial distress. Relative to nonviolators, violators are reliably more leveraged and have lower liquidity, profitability, and stock return performace. In year 0 (the year of technical violation), 85 percent of violators had losses, and 52 percent had losses in both year −1 and year 0. The corresponding percentages for nonviolators are reliably smaller at 24 and 13 percent. In addition, violators' mean stock return of −24 percent in the year prior to technical violation is significantly lower than the corresponding mean return for nonviolators (12 percent). These findings are interesting in conjunction with the evidence in table 2 that more than 90 percent of the technical violations occur on profitability, liquidity, and leverage covenants; they confirm that financial deterioration is a cause of technical violation.

Accounting policies, and changes therein, also have the potential to induce technical violation. To calibrate violators' choices against the population, we compare violators' portfolios of accounting policies—on depreciation, inventory, pension costs, and investment tax credits—to the portfolios of two random samples available from prior research. As shown in table 4, relative to the 300 firms in Zmijewski and Hagerman (1981) or the 83 firms in Press and Weintrop (1990), violators choose accounting policies that are significantly more income-increasing in both the year preceding and five years prior to technical violation. Although violators' actions are consistent with the positive theory prediction of choosing income-increasing accounting techniques to reduce the likelihood of violation (Watts and Zimmerman 1986), the evidence in year −5 shows that violators made such choices well before violation.

There are two implications to our accounting policy analysis. First, violators differ from nonviolators by using more liberal bases of accounting measurement, which tends to understate the differences on leverage, liquidity, and profitability reported in table 3.

---

[8] None of the nonviolators are in technical violation. However, it is possible that some of them are in debt service default and bankruptcy. Although we do not filter such firms out, if they are in the control sample, they make our comparisons more conservative.

Table 3

Selected Financial Characteristics of Sample Firms One Year Prior to Technical Violation Compared to Industry-Matched, Nonviolating Firms

Panel A. Selected Financial Characteristics:

	Sample Violators[a]				Nonviolators[b]				P-value,[c] Wilcoxon Test
Characteristic	*Mean*	*Median*	*Min.*	*Max.*	*Mean*	*Median*	*Min.*	*Max.*	
Market value of equity ($ million)	137.83	39.36	2.81	2,164.90	566.31	60.08	0.38	95,697.50	0.06
Total Assets ($ million)	493.47	104.72	7.60	9,573.50	1,002.93	90.33	0.18	63,832.72	0.10
Sales ($ million)	534.58	173.10	4.72	5,193.00	1,100.08	93.96	0.75	101,638.00	0.01
Total debt to total assets (%)	62.10	62.20	10.10	98.50	45.18	44.01	0.00	178.69	0.01
Long-term debt to total assets (%)	31.64	27.78	0.00	75.15	18.49	15.08	0.00	130.72	0.00
Current Ratio	2.18	1.91	0.23	14.21	2.62	2.13	0.10	18.13	0.00
Quick Ratio	1.35	1.00	0.12	13.76	1.68	1.21	0.04	102.81	0.00
Return on Assets (%)	−4.54	−1.99	−43.67	47.98	3.13	4.25	−29.78	46.27	0.00
Common stock return in the year prior to violation	−23.72	−28.89	−69.81	65.70	11.97	2.26	−84.36	196.07	0.00

Continued

Table 3—*Continued*

Panel B. Earnings Performance Prior to Technical Violation:

	Sample Violators	*Nonviolators*	P-*value,* *Wilcoxon Test*
Percentage of firms with losses:			
In year of technical default (year 0)	85.4%	24.5%	0.00
In years 0 and −1	51.6%	12.7%	0.00
In years 0, −1, and −2	25.3%	6.7%	0.00

[a] Data for 82 of the 91 violators are available on COMPUSTAT. For the remaining nine firms, the data are from annual reports and Moody's Industrial manuals.

[b] The nonviolator sample consists of all nonviolating firms in the same two-digit SIC industry as each technical violator, matched in the year of violator's default. Data are drawn from the COMPUSTAT primary, supplementary and tertiary annual files, and full coverage files. There are 34 different industries and a maximum of 5,698 firm-year observations available on any characteristic.

[c] The Wilcoxon two-sample test compares the distribution of sample firms' characteristics to the corresponding distribution for nonviolating firms. The reported P-values indicate the smallest probability of incorrectly rejecting the null hypothesis of no difference between the two samples.

Table 4

Accounting Strategies of Technical Violators in Comparison with Samples in Zmijewski and Hagerman (1981) and Press and Weintrop (1990)

Strategy Classification	*Sample Violators, Year −1 (Firm Data From 1982–1986)*		*Zmijewski & Hagerman (1981) (Firm Data From 1975)*		*Press & Weintrop (1990) (Firm Data From 1985)*		*Sample Violators, Year −5 (Firm Data From 1978–1982)*	
	# Firms	*Percentage*	*# Firms*	*Percentage*	*# Firms*	*Percentage*	*# Firms*	*Percentage*
1 (income-decreasing)	0	0.0	10	3.3	1	1.2	0	0.0
2	0	0.0	49	16.3	4	4.8	1	1.1
3	1	1.1	107	35.7	31	37.4	6	6.6
4	10	11.0	107	35.7	42	50.6	17	18.7
5 (income-increasing)	80	87.9	27	9.0	5	6.0	67	73.6
Totals	91	100.0	300	100.0	83	100.0	91	100.0

Notes: Because of limited availability and reliability of accounting policy data on COMPUSTAT, we calibrate the technical violators' policies against samples used in prior research. Strategy distributions for these samples are drawn from Press and Weintrop (1990, table 6). The strategy classifications are based on Zmijewski and Hagerman (1981), beginning with all four accounting choices (depreciation method, inventory, amortization of past pension service costs, and investment credit treatment) being income-decreasing and ending with all choices being income-increasing. Using the notes to each violator's financial statements in the year prior to violation (year −1), we consider an accounting choice income-decreasing if we find evidence of accelerated depreciation, LIFO, amortization of past pension service costs in less than 30 years, or deferral of investment tax credit. After 1985, the investment tax credit was abolished; after 1986, the treatment of pensions was changed by SFAS No. 87. For the 28 firms in violation in 1986 and 1987 period, we used the last annual report in which all policies were reported in conformance with Zmijewski and Hagerman (1981) and Press and Weintrop (1990). The mean strategy for the sample violators is 4.87 in year −1, 4.65 in year −5; the mean is 3.31 for the Zmijewski and Hagerman sample, 3.55 for the Press and Weintrop sample. (The mean strategy is obtained by multiplying the frequency of observed strategy by the income strategy number and dividing this product by the size of the particular sample.) A *p*-value of zero was obtained for the Z-score from Wilcoxon two-sample tests of the ranked strategy score of the sample of violators in year −1 versus that of the Zmijewski and Hagerman sample and that of the Press and Weintrop sample.

Second, if any voluntary change occurs in the year of violation, it is likely to be income-decreasing and *could* lead a firm into technical violation.

We thus investigate whether accounting changes in year 0 force firms into technical violation by negatively affecting balance sheets and income statements. We identify eight changes by reading auditors' reports and financial statement footnotes.[9] We find that technical violators implement three mandated changes in year 0: two were adoptions of SFAS No. 87, and there was one adoption of SFAS No. 96. In every case, the effect of the mandated change was to increase earnings per share, making it improbable that mandated changes led to technical violation.[10] In previous research on the effects of mandated accounting changes, only Frost and Bernard (1989) report instances of technical violation. However, of the two cases they report, one firm had been in violation because of financial distress in the year prior to the mandated change. Thus, we conclude that technical violations are induced by financial distress rather than accounting changes.

III. Method

We investigate the costs of technical violation in two ways. We identify changes in contracting terms imposed by lenders subsequent to violation, and then record the changes in management investment and financing decisions associated with lenders' requirements for partial or full repayment of the violated debt. This approach enables us to estimate refinancing costs by assessing changes in borrowing rates, and to investigate restructuring costs by examining management's program for debt repayment. We use data from two sources: (1) the actual debt agreements filed with the SEC and (2) management discussions and analyses (MDAs) and debt footnotes in Forms 10-K and annual reports in the year of violation and the subsequent year.

Information in Debt Agreements

Changes in debt agreement terms are assessed by comparing covenants in the debt agreements before and after violation. We refer to the debt agreement containing the violated covenant(s) as the violated agreement, and to the post-violation debt agreement supplanting the violated agreement as the renegotiated agreement.

We define renegotiation as the process of revising the terms of lending agreements. We determine that renegotiation occurs when financial statements explicitly discuss the renegotiation, or when we observe a modification of contract terms in a post-violation agreement. There are 61 firms that renegotiate with existing lenders or obtain new lenders.[11] We observe no evidence of renegotiation within one year following violation for the remaining 30 firms. Our assessment of changes in contract terms applies to the 61 firms that renegotiate.

---

[9] SAS No. 1, sec. 150.02, requires that an auditor "identify those circumstances in which . . . [accounting] principles have not been consistently observed in the current period in relation to the preceding period" (AICPA 1989). Accordingly, the auditor's report reflects material changes in accounting policies.

[10] The other five changes are voluntary and unlikely to have induced technical violation. Two changes from FIFO to LIFO decreased per share earnings by $.06 and $.08, and resulted in tax refunds of $757,000 and $1,500,000 from loss carrybacks. A switch to LIFO, which has a positive cash effect, is consistent with a distressed firm's using the change to stave off a default on debt service. Although these changes have a negative impact on earnings (average decrease of 5.8 percent), the firms were already in loss settings, thus making it unlikely that the changes led to technical violation. The other three changes either increased or had no material effect on net income.

[11] The renegotiation process is relatively fast. The mean time to complete renegotiation following violation is two months. This contrasts with the 15-month average required for troubled-debt restructurings described in Gilson et al. (1990, table 5).

We obtained violated and renegotiated agreements for 43 of the 61 firms by using the exhibit indices in item 14 of Forms 10-K in years −1, 0, and +1 to identify publicly available debt agreements. Since one of the 43 firms violates two agreements, we obtained 88 contracts (44 pre- and 44 post-violation). To ensure that the changes in contract terms we observe are based on the set of covenants in effect at the time of violation, we also acquired eight amendments to debt agreements made prior to violation. If the amendments were made in anticipation of violation, our contract terms comparisons are conservative. We examined a total of 96 contracts and related amendments.[12]

For the 18 remaining firms (out of 61) that renegotiated, we collected information about changes in contract terms by using MDAs and debt footnotes in years −1, 0, and +1. Debt agreements for these firms are not publicly available because of the relative size of the loan amount. Under SEC regulations S-K, item 601.b(4)(iii), firms are exempted from filing contracts with the SEC if the amount of the contract is less than 10 percent of total assets.

Information in Financial Statements

MDAs and debt footnotes in Forms 10-K and annual reports in the year of and year after violation are available for all sample firms. We used this information to corroborate changes in contracting terms when debt agreements are available and to assess changes when they are not. In addition, by referring to Forms 10-K and annual reports in the year following violation, we can identify changes in firms' investment and financing decisions associated with the violation.

IV. Empirical Findings

We discuss our empirical findings in four parts. First, we present evidence that there are incremental costs of financing associated with technical violation. Second, we show that lenders increase their monitoring of violators by imposing various new investing and financing covenants subsequent to violation. Third, we assess the restructuring costs imposed by lenders' demands for repayment. Finally, we show that the effects of technical violation differ according to lender response.

Refinancing Costs

Although changes in borrowing rates are an important, and common, consequence of technical violation, we can identify changes in borrowing rates for only 48 firms because of limitations in the availability of contracts and the extent of interest rate disclosures in financial statements. Information about interest rate changes is drawn exclusively from contracts for 43 of the 48 firms. For the remaining five firms, this information is collected from MDAs and debt footnotes. Rates increased for 31 of 48 firms, decreased for three firms, and were unchanged for 14 firms. As panel A of table 5 shows, the mean and median changes are increases of 80 and 55 basis points over prime, which is consistent with increased borrowing costs following violation.

---

[12] Of these 96 contracts, 42 were found in Forms 10-K, 18 in Forms 10-Q, 27 in Forms 8-K, and nine in registration statements. We purchased debt agreements when they were filed in 10-Qs, 8-Ks, or registration statements. We also purchased Forms 10-K from Disclosure, Inc., when those in the university collections available to us were missing or contained only an abridged shareholder version.

Table 5
Interest Costs Imposed by Technical Violation and Their Relation to Firm Characteristics

Panel A. Incremental Interest Costs ($n=48$):[a]

	Mean	*Median*	*Min.*	*Max.*
Change in contract interest rate (basis points relative to prime)	80.0	55.0	−100.0	500.0
Incremental interest costs for one year[b] ($000)	293.8	85.3	−30.9	2,224.1
Incremental interest cost/ market value of equity[c]	0.95%	0.40%	−0.3%	9.75%

Panel B. Relation of Incremental Interest Costs to Firm Characteristics ($n=48$):

Model: $INT/MVE_t = a_0 + a_1 LEV_t + a_2 ROA_t + a_3 RET_t + a_4 SEC_t + \epsilon_t$[d]

Coefficient Estimates (*t-statistics*)

Specification	a_0	a_1	a_2	a_3	a_4	Regression F (*p-value*)	*Adj.* R^2
I	−0.0027	0.0023	0.0065	−0.0113	0.0022	3.63	0.177
	(−0.51)	(2.92)*	(0.56)	(−1.23)	(0.43)	(0.01)	
II	−0.0014	0.0023	0.0071	−0.0113	—	4.86	0.191
	(−0.33)	(3.06)*	(0.62)	(−1.25)	—	(0.01)	
III	−0.0001	0.0026	0.0048	—	—	6.44	0.182
	(−0.01)	(3.56)*	(0.43)	—	—	(0.01)	
IV	−0.009	0.0026				12.92	0.196
	(−0.25)	(3.60)*				(0.00)	

[a] Changes in interest rates are determined by reference to the violated and renegotiated contracts and to MDAs and debt footnotes in years −1 to +1. For the 44 firms whose contract interest rates are variable rates based on the prime rate plus basic points, we observe 31 rate increases, three decreases, and 14 rates that are unchanged. The four firms with fixed rate contracts reflect one rate increase and three are unchanged.

[b] Measuring the effect of interest rate changes on cash flow required knowledge of the amount outstanding at every point in time in the period subsequent to violation. Since such data are not available, we base our calculations on the amount of violated debt outstanding at the end of year +1 relative to the year of violation. Thus, we implicitly assume that the amount outstanding at year-end approximates the average amount outstanding. This assumption is reasonable if inadequate cash flows prevent violators from reducing their loans.

[c] Market value of equity prior to violation is computed as price on day −2 of violation announcement times number of shares outstanding. The data are from the CRSP daily returns tape.

[d] Variable definitions:

INT/MVE = incremental interest cost/market value of equity;
LEV = ratio of total debt to the market value of equity at the end of year 0;
ROA = net income to total assets in year 0;
RET = return on common stock for fiscal year 0; and
SEC = 1 if violated debt agreement is secured, and 0 otherwise.

* Significant at the 5 percent level.

We compute incremental interest cost for firm j as the product of the post-violation change in the rate of borrowing (Δi_j) times the expected amount of outstanding borrowing ($LOAN_j$):

$$\text{Incremental interest cost}_j = \Delta i_j * LOAN_j. \tag{1}$$

Interest rate changes, Δi_j, are measured as the post-violation change in basis points over prime for the 44 firms with variable rate debt agreements. For the four firms with fixed rate contracts, there was one rate increase, and three were unchanged. For all loans, we assume that the expected amount outstanding is constant in the year subsequent to violation and is approximated by the amount outstanding at the end of fiscal year +1. This assumption is reasonable if inadequate cash flows prevent violators from substantially reducing their loans.

We estimate the average one-year incremental interest cost of violation to be $293,800. This cost represents 0.95 percent of the market value of equity, measured two days prior to the first announcement of technical violation. For a two-year period, the incremental interest cost represents 1.86 percent of market value of equity.[13] With the assumption that the next negotiation occurs within two years and that the effects of increased interest rates last until the next negotiation, actual refinancing costs arising from interest rate changes fall between our one- and two-year estimates.

These assumptions are plausible for several reasons. First, the one-year estimate is the lower bound of incremental interest costs since we find the interest rate changes still in effect one fiscal year after violation. Second, we expect the lending agreements to be renegotiated in the second year, given evidence in Frost and Bernard (1989) that private debt agreements (of which our sample is mostly comprised) are negotiated on average every 15 months and given that the mean time between successive negotiations is 21 months in our sample. Third, the interest rate set at the next negotiation depends on how the firm—under the increased scrutiny of lenders—has performed post-violation; rates could increase or decrease. The outcome of subsequent negotiations is dependent on post-violation developments, and we do not attribute their effects to the instance of technical violation we study.

These findings are evidence of direct refinancing costs imposed by technical violation. Since previous researchers have used surrogates to represent these costs, we examine the relation between incremental interest costs and several firm characteristics: leverage, profitability, stock return in the year prior to violation, and provision of collateral. In table 5, panel B, we report estimates of the following regression:[14]

$$INT/MVE_i = a_0 + a_1 LEV_i + a_2 ROA_i + a_3 RET_i + a_4 SEC_i + \epsilon_i, \tag{2}$$

---

[13] Incremental interest costs are discounted at the risk-free rate under the assumption that debt must be paid when equity holders' residual claims are positive. We treat the cash flows as if they were riskless. The magnitude of incremental interest costs for two years is similar (1.72 percent of the market value of equity) when we use the risky equity discount rate.

[14] We obtain similar results using total debt to total assets as a proxy for leverage. However, the relation is much weaker when measuring leverage as the ratio of long-term debt to total assets, probably because of the high incidence of long-term debt reclassifications for technical violators. We also considered two other variables: size and trends in profitability. These two variables did not enhance the specification and are not reported.

where,

INT/MVE = incremental interest cost/market value of equity;
LEV = ratio of total debt to the market value of equity at the end of year 0;
ROA = net income to total assets in year 0;
RET = return on common stock for fiscal year 0; and
$SEC = 1$ if violated debt agreement is secured, and 0 otherwise.

The distinctive feature in all regressions in table 5 is that only leverage is significantly correlated with incremental interest costs. This is an important finding because previous researchers have used leverage to proxy for costs associated with debt covenant violations, and we show that the magnitude of refinancing costs is positively associated with leverage. A word of caution is in order: the adjusted R^2 is only 19.6 percent, which suggests that leverage is a noisy proxy for incremental interest costs.

Because interest rates are available for only 48 firms, we estimate the magnitude of incremental interest costs for the entire sample by extrapolating our findings in two ways. First, with the assumption that interest rates are unchanged for the 43 firms on which data are unavailable, the average one-year (two-year) incremental interest cost for the whole sample represents 0.50 percent (0.98 percent) of market value of equity. Alternatively, we assume that interest rate changes are similar for the 43 firms with missing data and estimate their interest costs using leverage (specification IV in table 5).[15] Using this procedure, we estimate that the average one-year (two-year) incremental interest costs for the entire sample of 91 technical violators is 0.84 percent (1.63 percent) of the market value of equity. Under either assumption, technical violations impose significant refinancing costs on firms.

Covenant Changes Subsequent to Violation

Changes in constraint boundaries and new accounting, investing, and financing covenants are another consequence of technical violation. In table 6, panel A, we compare the levels of five accounting covenants studied in prior research (Press and Weintrop 1990) for which we can identify specific constraint boundaries by reading violated debt agreements pre- and post-violation. In total, there are 163 observations across the five dimensions, which includes 74 of the 126 violated accounting covenants reported in table 2.[16] In total, 56 constraints (including 42 of the 74) are relaxed and 107 are tightened or left unchanged. Boundaries on net worth, current ratio and working capital, and leverage are reduced to accommodate violators' financial positions. These covenants are the ones most commonly violated (see table 2), and relaxation often sets a more attainable benchmark for the future, without removing the violation.[17] Interest-

---

[15] Applying specification IV seems reasonable because of similarities between the 48 firms for which we have interest data and the 43-firm complement that has missing observations. Using Wilcoxon two-sample tests, we find that both subsamples are similar in size (market value of equity), leverage (total debt to market value of equity), stock return in the year prior to violation, and the existence of loan security. They differ at the 5 percent level of significance in that the 48-firm subsample is less profitable (return on assets).

[16] We could not track 52 constraint levels post-violation. In table 2, we require only pre-violation data to determine whether a constraint is violated; in contrast, table 6 requires that specific constraint boundaries be available pre- and post-violation.

[17] For example, the constraint on tangible net worth for Dravo Corporation was changed from $175 million to $130 million following its 1987 loss of $126 million. Despite the relaxation, the firm remained in violation and required a waiver because its net worth at December 31, 1987, was only $90 million.

Table 6
Changes in Debt Agreement Covenants Subsequent to Violation

Panel A. Percentage Change in the Level of Existing Accounting Covenants:[a]

	Number of Firms[b]	Mean	Median	Min.	Max.
Leverage	33	25.5	0.00	−89.00	510.00
Tangible net worth	40	−4.66	−15.85	−84.28	300.00
Working capital	21	−22.49	−20.84	−90.50	85.00
Current ratio	24	−13.42	0.00	−58.33	8.00
Restricted retained earnings	45	43.49	0.00	−66.00	900.00
(Total Observations = 163)					

Panel B. Additional Covenants in Renegotiated Agreements:

Accounting Covenants	
No consecutive quarterly losses	10
Fixed charge coverage	6
Investing Covenants	
Limitation on capital expenditures	27
Prohibition against stock repurchase	15
Prohibition of acquisition or merger	11
Sale of non-current assets to repay debt	8
Lender approval for sale of assets	6
No research and development joint ventures	2
Financing Covenants	
Collateral increases	60
Prohibition of additional debt	12
Creditor approval of debt refinancing	11
No sale-lease back	8
Additional compensating balances	8
(Total Additional Covenants = 184)[c]	

Continued

ingly, increases in the restrictiveness of accounting constraints occur on dividends, which are not the source of violation in our sample (cf. table 2), but this is consistent with lenders' safeguarding the asset base to prevent further dissipation of their claims.

Along with changes in levels of accounting-based covenants, renegotiated agreements contain numerous new investing and financing covenants that increase lenders' monitoring of firms' activities. We tabulate these accounting, investing, and financing covenants in table 6, panel B. Interestingly, there are few additional accounting-based covenants (16 of 184 or 9 percent). Apparently, only slight adjustments to accounting-based monitoring in debt agreements are required. However, lenders increase their control over violators' operations by restricting investing and financing activities. Investing opportunities are limited by new covenants on capital expenditures, stock repurchases, and acquisitions. Most firms are required to provide greater loan collat-

Table 6—*Continued*

Panel C. Relation of Change in Number of Covenants Subsequent to Violation and Firm Characteristics:

Model: $ADDCON_i = a_0 + a_1 LEV_i + a_2 ROA_i + a_3 RET_i + a_4 SEC_i + \epsilon_i$[d]

Coefficient Estimates (t-statistics)

Specification	a_0	a_1	a_2	a_3	a_4	*Regression* F (p-*value*)	*Adj.* R^2
I	0.2224	−0.0088	−0.5931	−0.2687	0.0279	2.10	0.067
	(1.95)**	(−0.52)	(−2.24)*	(−1.26)	(0.25)	(0.09)	
II	0.2394	−0.0083	−0.5846	−0.2716	—	2.83	0.082
	(2.67)*	(−0.51)	(−2.25)*	(−1.29)	—	(0.05)	
III	0.2748	−0.0019	−0.6615	—	—	3.38	0.073
	(3.20)*	(−0.13)	(−2.60)*	—	—	(0.04)	

[a] The percentage change in level of the constraint is the change in the constraint boundary divided by the pre-violation boundary. For leverage, an increase in the level indicates that the firm can take on more debt, and thus represents a relaxation. For the other four covenants, a positive (negative) change indicates tightening (relaxation).

[b] The number of firms represents a subset of the 61 technical violators that renegotiate contracts following violation. Each subset consists of the firms for which we can identify the levels of a given constraint pre- and post-violation.

[c] This represents an average of three new covenants per firm (184/61), a 27 percent increase in the number of covenants.

[d] Variable definitions:

ADDCON = percentage change in number of covenants;
LEV = ratio of total debt to the market value of equity at the end of year 0;
ROA = net income to total assets in year 0;
RET = return on common stock for fiscal year 0; and
SEC = 1 if violated debt agreement is secured, and 0 otherwise.

eral, and creditors increase control of financing choices by limiting additional debt and requiring prior approval for new financing.

As table 6, panel B shows, renegotiated agreements add an average of three new covenants, a 27 percent increase in the number of covenants. Employing a model similar to equation (2), we examine the relation between the percentage increase in number of covenants (*ADDCON*) and firm characteristics. OLS estimates of the following regression appear in table 6, panel C:

$$ADDCON_i = a_0 + a_1 LEV_i + a_2 ROA_i + a_3 RET_i + a_4 SEC_i + \epsilon_i. \qquad (3)$$

Our specification of the dependent variable, *ADDCON*, assumes there are no qualitative differences in the effect of new covenants; the explanatory variables are as defined for equation (2). In all specifications of equation (3) in table 6, leverage is unrelated to the percentage change in the number of contraints, but profitability is significantly correlated. If the percentage increase in number of covenants proxies for some of the costs of increased lender control, the results indicate that neither leverage nor *ROA* (given the low adjusted R^2 of 7.3 percent) is a good surrogate for these costs.

Restructuring Costs

Within one year of violation, violators change their investing and financing decisions to conserve or improve cash flows. Among the 91 firms, we find that violators divest assets (54 firms), reduce or omit dividends (12), reduce research and development expenditures (12), privately place common stock with major shareholders (nine), issue warrants or refinance and exchange debt (eight), cut their labor forces (six), and terminate pension plans (five). These changes are made both to compensate for poor performance (85 percent of violators incur losses in the year of violation) and to comply with lenders' demands for repayment.

Although we cannot distinguish management actions designed to comply with lenders' demands from those intended to improve performance, we can identify a subset of actions that financial statements describe as arising from technical violation. These consist of debt refinancing and asset divestiture to repay violated debt. In eight cases, firms either issued warrants to lenders or refinanced or exchanged debt to repay the violated debt. Of 54 firms divesting subsequent to violation, 42 state that the proceeds were used to repay the violated debt. We view this subset of actions as firms' responses to lenders' demands for partial or full acceleration, and their effects as a source of restructuring costs.

Warrants and Debt Refinancing Impact on Cash Flow. Eight firms issued warrants to lenders as required in the renegotiated agreement or issued a new debt and stock package in exchange for existing (old) debt. When sufficient data are available, the following three costs are estimated for each of these firms.[18]

1. The cost of issuing warrants to lenders, which is equal to the market value of warrants minus any consideration received. The market value of warrants issued is computed by using Black-Scholes option pricing as applied to warrants in Stoll and Whaley (1992).[19]
2. The transaction cost of issuing new securities in exchange for debt (investment banker and legal fees). We identify this cost when reported in the long-term debt and shareholders' equity footnotes to financial statements.
3. The effect on cash flows of any difference in interest rate between the old and new debt, after taking into account changes in the general level of interest rates. The change in interest cost is computed as the difference between the amount of old debt times its effective interest rate minus the amount of new debt issued times its effective interest rate. The change in interest cost is discounted for the period remaining before maturity of the old or new debt, whichever is shorter.

The average cost of warrants issued to lenders for four firms is $711,600, or 1.03 percent of market value of equity prior to violation (MVE). The average issuance cost for five firms is $471,800, or 1.49 percent of MVE. The average incremental interest cost for three firms is $1,102,000, or 7.52 percent of MVE.[20] For the eight firms, the average (median) total cost represents a market value loss of 4.26 percent (1.11 percent).

---

[18] The eight firms are Anacomp, Artra, Berkey, Centronics, Custom Energy Services, Digicon, Geothermal Resources, and Publicker Industries. A synopsis of the refinancing packages these firms offered is available from the authors.

[19] The market value of warrants is estimated by using R. Whaley's option valuation program (OPTVAL). The data used in the estimation are available from the authors upon request.

[20] Because rates on exchanged debt are fixed, we adjust for changes in the level of interest rates in the economy by using changes in Moody's Aaa corporate bond yield between the year of issuance of the old debt and the new debt. For the three firms for which we estimate incremental refinancing costs, we increase the effective interest rate on the new debt by 125 to 226 basis points, the decrease in Moody's Aaa corporate bond yields. Moody's data are reported in Federal Reserve *Bulletins* (1984–1986).

Extrapolating to the 91 firm sample, we obtain an average cost of 0.37 percent. Combining these debt exchange costs with our previous estimates of incremental borrowing costs yields estimates of financing-related costs from technical violation of between 1.21 percent and 2 percent of MVE.[21]

Effect of Divestiture on Firm Performance. Forty-two sample firms divest subsequent to violation and report that the proceeds repaid violated debt. For eight of these firms, the renegotiated contracts require divestiture; for the remaining 34 firms, disposal of noncurrent assets to repay the violated debt is discussed in MDAs and annual report footnotes in year +1. Divestiture imposes costs if violators abandon or sell profitable projects to repay debt. Alternatively, divestiture could be efficient if firms discontinue unprofitable projects. In addition, if technical violation is a consequence of poor management, the reduction in the asset base reduces managers' opportunities to shirk, consume perquisites, and dissipate resources.

To determine the consequences of divestiture for sample firms, we compare the performance of 42 divesting violators to 49 nondivesting violators. We pair each violator to a specific control firm and adjust for industry and size by computing paired differences between violators and their controls. We then evaluate the paired differences across subsamples of divesting and nondivesting violators.[22]

Table 7 reports the pre- and post-violation comparisons. Before technical violation, there are no differences in size, profitability, and capital expenditures betweeen the divesting and nondivesting violators. However, following violation, divesting firms are both significantly smaller and less profitable. The size result is not surprising since by definition divestiture reduces the asset base, but the performance results are consistent with abandonment of profitable investment projects post-violation. A *caveat* is needed: if poor performers self-select into the divesting group, our matching procedure may not overcome this selection bias. Given this caution, the evidence suggests that lenders' demands for repayment impose costs on divesting violators that arise from lost returns on investments, although we cannot estimate their magnitude.

In summary, our investigation of the effects of technical violation produces the following evidence. First, lenders exercise options to change contract terms. Prior research alludes to such actions, but we document that lenders increase interest rates, increase their control of violators' operations through additional covenants, and sometimes demand partial or full repayment. Second, we show that lenders' actions adversely affect the value of violating firms by imposing refinancing and restructuring costs. Thus, we provide direct evidence that technical violation is costly. We have argued that the effects of technical violation result from lenders' exercise of various contractual rights, and we now extend our analysis to examine whether differential effects arise conditional on lender response.

Consequences of Technical Violation Conditioned by Lender Response

We examine whether the effects of technical violation vary according to lender response by partitioning the sample on the basis of whether lenders grant waivers. We

---

[21] Estimating equation (2) with all these financing-related costs as the dependent variable produces results similar to those in table 5.

[22] Each control firm is chosen from the population of COMPUSTAT firms according to two criteria: (1) the control firm is in the same four-digit industry as the violator and (2) within the industry, the control firm is the closest in size (total assets) to the violating firm. A Wilcoxon two-sample test fails to reject the null hypothesis that violator and control firms' sizes are drawn from the same population.

Table 7
Lender Repayment Demands, Asset Divestiture, and Firm Performance Subsequent to Technical Violation

	Pre-Violation Median Deviation From Control Firm (Years −3 to −1)			*Post-Violation Median Deviation From Control Firm (Years +1, +3)*		
Attributes	*Violator Divesting To Repay vs. Control (N=42)*	*Violator Not Divesting vs. Control (N=49)*	*Wilcoxon Z*	*Violator Divesting To Repay vs. Control (N=42)*	*Violator Not Divesting vs. Control (N=49)*	*Wilcoxon Z*
Levels						
Sales ($MM)	−1.8	+10.1	0.44	−65.2	−32.9	−1.91**
Market value of equity ($MM)	−99.8	−85.4	0.23	−148.1	−89.9	−1.45
Number of employees	+250.0	+828.0	0.40	−448.0	+383.0	−1.67**
Ratios						
ROA	−12.79%	−9.67%	0.33	−17.71%	2.49%	−2.15*
Profit margin	−6.78%	−13.28%	−0.12	−20.16%	4.66%	−2.31*
Capital expenditure to fixed assets	−3.65%	−9.17%	−0.13	−15.08%	−14.99%	0.09

Notes: Forty-two out of 91 firms divest subsequent to violation in order to repay all or part of the violated debt. For eight of these, the renegotiated contract requires the disposal of noncurrent assets to repay or reduce debt. For the remaining 34 firms, the disposal for repayment is described in the MDAs or in the debt footnotes in years 0 and +1. Control firms are chosen from COMPUSTAT and have the same four-digit industry as the violator and are closest in size (total assets) to the violating firm three years prior to violation. Wilcoxon two-sample tests fail to reject the null hypothesis that violator and control firms' sizes are drawn from the same population three years prior to violation. Let A_{vi} and A_{ci} be the attribute *i* values for violating firm *v* and its control *c*. Deviation from control for violating firm *i* is computed as the difference $A_{vi}-A_{ci}$. Data definitions: Sales and number of employees are measured at the end of firms' fiscal years. Market value of equity is the product of price and shares outstanding at the end of firms' fiscal years. ROA is the return on assets computed as the ratio of income before extraordinary items to total assets. Profit margin is the ratio of income before extraordinary items to sales. Capital expenditures are deflated by net fixed assets. All data are measured by using the average for the three-year periods −3 to −1 and +1 to +3.

* Significant at the 5 percent level.
** Significant at the 10 percent level.

define a waiver as a suspension of contractual rights for a specified period of time.[23] Since lenders can extract fees and concessions in exchange for waivers, renegotiation and waiver are not independent outcomes. Thus, we partition the 91 violators into four subsamples, A through D, as follows:

---

[23] Although the duration of a waiver is a pertinent component of lender response, unavailable data preclude a partition based on duration. We thus use a 1/0 partition to distinguish the 65 violators that receive waivers from the 26 firms that do not. We compare the groups along dimensions of size, profitability, leverage, and loan security. We find that waived firms are larger (in terms of total assets and market value of equity) and less leveraged (in terms of total debt to either total assets or market value of equity). It is, however, outside the scope of this study to assess the factors affecting lenders' decisions to grant waivers.

Subsample A:

Waiver following renegotiation: Fifty-three violators successfully renegotiate with existing lenders and obtain a waiver between years 0 and +1 by making concessions to lenders.[24]

Subsample B:

Waiver but no renegotiation: Twelve firms are granted waivers without any alteration of contract terms.

Subsample C:

No waiver following renegotiation: Eight firms fail to arrive at an agreement with existing lenders, and the violations remain unwaived. These eight firms obtain financing from a new lender.

Subsample D:

No waiver and no renegotiation: Eighteen firms neither renegotiate nor obtain a waiver within one year of technical violation. It is possible that lenders are slow to act with these firms, fearing the repercussions of bad faith allegations in lender liability suits (Johnson and Gaffney 1988; West and Haggerty 1988).[25] Alternatively, debt agreements might automatically provide for penalties in an event of default.

We assess whether the effects of technical violation differ across the subsamples by comparing the values of four measures presented previously: post-violation changes in interest rates, incremental interest costs, the number of added covenants, and the number of firms divesting to repay debt. We also report market model estimates of stock price reactions to announcements of technical violation.[26]

---

[24] Renegotiation modifies contract terms but does not necessarily take the firm out of violation. If renegotiation tightens or leaves accounting constraint boundaries unchanged, a waiver is necessary. Even when accounting constraints are relaxed, this often merely sets a more attainable benchmark for the future and a current waiver is required. For example, Fischbach Corporation's lender relaxed the requirement for tangible net worth and working capital minima from $142 and $110 million to $115 and $37 million, respectively, following the firm's 1987 violation. Fischbach was then in compliance with the working capital constraint, but was still $13 million deficient in net worth. Therefore the lender waived the violation. Custom Energy Services is another example. Both leverage and tangible net worth covenants were violated and relaxed, but the firm's debt level still exceeded the maximum renegotiated leverage level.

[25] Since the early 1980s, lenders have lost a significant number of suits brought by disgruntled borrowers, and substantial judgments have been awarded when lenders refused to make loan advances (e.g., *K.M.C. Co., Inc. v. Irving Trust Co.*), used loan covenants improperly (e.g., *State National Bank of El Paso v. Farah Manufacturing Co., Inc.*), and improperly accelerated repayment (e.g., *Sahadi v. Continental Ilinois National Bank & Trust Co.*). The commercial-lending literature from the 1980s is replete with advice about avoiding lender liability suits. Weissman (1986) and Johnson and Gaffney (1988) discuss the marked increase in lender liability litigation.

[26] We view the stock price results as indirect and suggestive evidence because announcements of violation affect stock prices only if violation imposes costs on a firm or if investors view the announcement itself as a signal of lower future cash flows from continuing operations. We do not attempt here to distinguish between these effects. Announcements of violation are often contaminated by concurrent announcements of earnings, audit qualifications, reclassification of long-term debt to current liabilities, operation restructuring, and waivers of violation; we do not control for their confounding effects. Finally, we record consequences of violation on the basis of information available subsequent to the violation announcement. For example, our waiver criterion is based on receipt of a waiver by the end of the fiscal year following violation. We address these complications in another manuscript (Beneish and Press 1992).

Table 8
Consequences of Technical Violation Conditional on Lender Response

Panel A. Consequences of Technical Violation Across Subsamples:

Attributes	*A: Waiver Following Renegotiation (N=53)*	*B: Waiver But No Renegotiation (N=12)*	*C: No Waiver Following Renegotiation (N=8)*	*D: No Waiver and No Renegotiation (N=18)*
Mean change in interest rate (basis points over prime)	84.0	0.0	140.0	33.0
Mean (incremental interest cost/MVE)	0.81%	0.0	2.35%	0.54%
Mean number of additional covenants	2.64	0.0	5.25	0.0
Percentage of firms divesting to repay debt	47.1%	0.0%	50.0%	72.2%
CAPE(−1,+1)[a]	−2.76%	0.23%	−10.01%	−5.17%

Continued

Statistics about the values of these attributes in each subsample are presented in panel A of table 8. Note that, for firms that do not renegotiate the violated agreements, we are able to observe divestitures and stock price effects but no changes in contract terms. An exception occurs for three firms in Subsample D; for these, the violated debt agreement contains provisions for an interest rate change in the event of default.

We compare Subsamples A and B to determine the effect of renegotiation on the cost of obtaining waivers. As shown in table 8, panel B, we can reject null hypotheses of no difference at the 5 percent level for four of five comparisons. Firms that obtain waivers following renegotiation (Subsample A) face higher interest rates and more covenants, divest more frequently, and have poorer stock price performance. Since renegotiation is often necessary to obtain a waiver, these findings suggest that waivers are not costless.

We compare Subsamples A and C to determine the effect of unsuccessful renegotiation. We find that the eight firms that obtain new lenders rather than waivers bear larger interest costs, have twice as many new covenants, and suffer greater stock price losses. This is consistent with higher costs of technical violation when lending agreements are terminated and violators seek a new lender. There appears to be significant differential costs to credit termination reltaive to successful renegotiation.

Finally, we compare Subsamples A and D. The 18 firms in Subsample D neither renegotiate nor obtain waivers within a year of violation. We find that these firms tend to divest more often to repay debt, and experience significantly greater stock price losses at the time of announcements of violation. Thus, it appears that unwaived violations have more adverse effects on violators' value.

The evidence suggests that waivers are costly for firms that receive them. Nonetheless, the costs of technical violation are lower for firms that can obtain a waiver than for those that cannot.

Table 8—*Continued*

Panel B. Comparisons Across Subsamples:

		t-statistic (p-value)		
Attributes	*Test*	*Subsample A vs. B*	*Subsample A vs. C*	*Subsample A vs. D*
Mean change in interest rate (basis points over prime)	Wilcoxon Z two-sample[b]	−1.86* (0.03)	1.44** (0.08)	−0.74 (0.25)
Mean (incremental interest costs/*MVE*)	Wilcoxon Z two-sample	−0.36 (0.71)	−1.60** (0.06)	−2.07 (0.94)
Mean number of additional constraints	Wilcoxon Z two-sample	−2.64* (0.01)	2.05* (0.02)	0.40 (0.36)
Percentage of firms divesting to repay debt	Binomial proportions[c]	6.83* (0.00)	0.15 (0.89)	−2.06* (0.02)
CAPE(−1,+1)[a]	Return difference *t*-statistic[d]	2.66* (0.00)	−6.35* (0.00)	−2.58* (0.02)

Notes: We define renegotiation as changes in the terms of debt contract between lender and violator. We define waiver as suspension of contract rights for a specified period of time. Renegotiation and waiver status are determined by reading *MDA*s and annual report debt footnotes in years 0 and +1 relative to the fiscal year of violation. We read year +1 to ascertain the outcome of renegotiation when such outcome could not be determined from year 0 data. The variables are defined in tables 5 and 6.

[a] The estimation of cumulative average prediction errors (*CAPE*) is determined as follows. Market model prediction errors PE_{it} are calculated as $PE_{it}=R_{it}-(\hat{\alpha}+\hat{\beta}_i R_{mt})$, where R_{it} and R_{mt} are the continuously compounded rates of return on the common stock of firm *i* and the equally weighted NYSE and ASE index on event day *t* (Fama 1976). Market model parameters are estimated over 300 trading days, from day +61 to +361 relative to the day of announcement of violation; $CAPE_k=\Sigma_{t=1}^{k} APE_t$, where $APE_t=1/N\Sigma_{i=1}^{N} PE_{it}$. *t*-statistics are calculated as follows: $t(CAPE_k)=CAPE_k/(ks^2(APE_t))^{1/2}$, where $s^2(APE_t)$ is the estimated variance of average prediction errors over days +61 to +160 relative to event day 0 and *k* is the number of days in the cumulation period. The statistics are distributed approximately *t* with 99 degrees of freedom. For the entire sample, the *CAPE* from days −1 to +1 is −3.52 percent, with a *t*-statistic of −3.89.

[b] Wilcoxon two-sample tests compare the distributions of various attributes across subsamples by waiver and renegotiation outcome. The statistics reported are based on one-tailed tests.

[c] Since the divesting behavior is coded as a 1, 0 dummy variable, we compare the percentage of firms divesting across subsamples using a test of binomial proportions.

[d] The difference in *CAPE*(−1,+1) between subsample *i* and *j* is computed as:

$$Diff_{ij}=\sum_{t=-1}^{+1}(APE_{it}-APE_{jt}),$$

where *APE* denotes the average prediction error and *i* and *j* two subsamples. The statistical significance of $Diff_{ij}$ is evaluated with the following statistic: $t(Diff_{ij})=Diff_{ij}/(3s_{ij}^2(APE))^{1/2}$ where $s_{ij}^2(APE)$, the variance of the average prediction errors of a portfolio containing all firms in subsamples *i* and *j*, is estimated over days +61 to +160 relative to day 0. The statistic is approximately distributed *t* with 99 degrees of freedom.

* Significant at the 5 percent level.
** Significant at the 10 percent level.

V. Conclusion

This study provides direct evidence about the costs facing firms that violate accounting covenants in debt agreements. We show that, as a result of technical violation, lenders increase their control of firm activities and exercise various contractual rights, which imposes refinancing and restructuring costs on violators. Further, we furnish evidence that the costs vary with lender response: violation is less costly for firms that can obtain a waiver than for those that cannot, but, more important, waivers are costly, since lenders extract fees and concessions in exchange for granting them.

In interpreting the findings in this study, one should note that the sample likely represents the upper end of the population of technical violators. Because current GAAP and SEC rules require disclosure only of material, uncured technical violations, it is impossible to assess how many violations occur relative to how many are reported.

This is one of the first studies to show that technical violation of accounting-based covenants is costly. The average costs we estimate range between 1.2 percent and 2 percent of market value of equity. Linking these costs to characteristics of violating firms, we show that leverage proxies for the magnitude of the costs imposed by technical violation.

References

American Bar Foundation. 1971. *Commentaries on Indentures.* Chicago: ABF.

American Institute of Certified Public Accountants. 1989. *AICPA Professional Standards,* Vol. 1, *U.S. Auditing Standards.* Chicago: Commerce Clearing House.

Ang, J., J. Chua, and J. McConnell. 1982. The administrative costs of corporate bankruptcy: A note. *Journal of Finance* 37 (March): 219–26.

Beneish, M. D., and E. Press. 1993. The resolution of technical default. Working paper, Duke University, Durham, NC.

Fama, E. 1976. *Foundations of Finance.* New York: Basic Books.

Financial Accounting Standards Board. 1983. *Statement of Financial Accounting Standard No. 78: Classification of Obligations That Are Callable by the Creditor: An Amendment of ARB No. 43, Chapter 3A.* Stamford, CT: FASB.

——. 1985. *Statement of Financial Accounting Standard No. 87: Employers' Accounting for Pension Plans.* Stamford, CT: FASB.

——, Emerging Issues Task Force. 1986. *Classification of Obligations When a Violation is Waived by the Creditor.* Stamford, CT: FASB.

Frost, C., and V. Bernard. 1989. The role of debt covenants in assessing the economic consequences of limiting capitalization of exploration costs. *The Accounting Review* 64 (October): 788–808.

Gilson, S. 1990. Bankruptcy, boards, banks, and blockholders: Evidence on changes in corporate ownership and control when firms default. *Journal of Financial Economics* 27 (October): 355–87.

——, L. Lang, and K. John. 1990. Bankruptcy, boards, banks, and bondholders: Evidence on changes in corporate ownership and control when firms default. *Journal of Financial Economics* 27 ( ): 315–53.

Holthausen, R. 1981. Evidence on the effect of bond covenants and management compensation contracts on the choice of accounting techniques: The case of the depreciation switch-back. *Journal of Accounting & Economics* 3 (March): 73–109.

Johnson, D., and T. Gaffney. 1988. Lender liability: Perspectives on risk and prevention. *Banking Law Journal* 105 (July-August): 325–48.

Leftwich, R. 1981. Evidence of the impact of mandatory changes in accounting principles on corporate loan agreements. *Journal of Accounting & Economics* 3 (March): 3–36.

Lys, T. 1984. Mandated accounting changes and debt covenants: The case of oil and gas accounting. *Journal of Accounting & Economics* 6 (April): 39–65.

Press, E. 1988. The role of accounting-based contractual provisions and financial statement effects in explaining a debt extinguishment: The case of in-substance defeasance. Unpublished Ph.D. dissertation, University of Oregon, Eugene.

——, and J. Weintrop. 1990. Accounting-based constraints in public and private debt agreements: Their association with leverage and impact on accounting choice. *Journal of Accounting & Economics* 12 (January): 65–95.

Securities and Exchange Commission. 1988. *SEC Accounting Rules*. Chicago: Commerce Clearing House.

Stoll, H., and R. Whaley. 1992. *Futures and Options: Theory and Applications*. Cincinnati, OH: South-Western.

Trombley, M. 1989. Accounting method choice in the software industry: Characteristics of firms electing early adoption of SFAS No. 86. *The Accounting Review* 64 (July): 529–38.

Warner, J. 1977. Bankruptcy costs: Some evidence. *Journal of Finance* 32 (July): 337–47.

Watts, R., and J. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice Hall.

Weiss, L. 1990. Bankruptcy resolution: Direct costs and violation of priority of claims. *Journal of Financial Economics* 27 (October): 285–314.

Weissman, M. 1986. Lender liability: The obligation to act in good faith and deal fairly. *Journal of Commercial Bank Lending* 69 (December): 2–14.

West, G., and M. Haggerty. 1988. The "demandable" note and the obligation of good faith. *Uniform Commercial Code Law Journal* 21 (Fall): 91–137.

Zmijewski, M., and R. Hagerman. 1981. An income strategy approach to the positive theory of accounting standard setting/choice. *Journal of Accounting & Economics* 3 (August): 129–49.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 258–272

Stock Market Effects of the Closeness to Debt Covenant Restrictions Resulting from Capitalization of Leases

Samir M. El-Gazzar
Pace University

SYNOPSIS AND INTRODUCTION: This article examines the association between lessees' market returns and their changes in the tightness of the debt covenant constraints resulting from compliance with SFAS No. 13 (1976). This paper differs from previous research in that it uses covenant-based measures. Similar to Hughes and Ricks (1984) and Schipper and Thompson (1983), it examines cross-sectional stock return dependencies that exist when industry effects cannot be randomized over time.

The paper begins by identifying a sample of lessees who retroactively capitalized leases as a result of SFAS No. 13. Actual debt contracts of sample firms are then analyzed to identify accounting-based restrictions, the default value of each restriction, and the definition of the accounting variables used in the covenant. Next, the percentage increase in the tightness of covenant restrictions arising from adopting SFAS No. 13 is calculated.[1] Finally, cross-sectional tests of the relationship between changes in covenant tightness and changes in security prices that accompanied the events leading to SFAS No. 13 are performed.

The analyses reveal several results. First, retroactive capitalization of off-balance sheet leases would have caused significant increases in the tightness of the debt covenant restrictions. Second, affected lessees experienced negative market returns contemporaneously with the disclosure of two of the events that led to the promulgation of SFAS No. 13. However,

---

[1] Descriptions of this calculation are presented in section II and in the appendix.

I am grateful for the valuable comments of Victor Pastena, Steven Lilien, Victor Bernard (associate editor for this article), and an anonymous referee. I also appreciate comments (on earlier drafts of this paper) from presentations made at Rutgers University, Baruch College, and the AAA annual meeting in Hawaii. The research assistance of Yuki Kim is appreciated. The Author is also associated with Tanta University in Egypt.

Submitted June 1989 and May 1991.
Accepted October 1992.

the magnitude of the reduction in market returns is correlated with the impact of SFAS No. 13 on the tightness of debt covenant restrictions.

Finally, there appear to be important differences in the structure of debt covenants. Private debt covenants have tighter financial restrictions, while public debt covenants have more nonaccounting-based provisions such as sinking fund, security, and seniority of the debt.

Key Words: *Lease accounting, Debt covenants, Economic effects.*

Data Availability: *A list of sample firms is available from the author. Data from debt contracts used in this study will be available for research purposes upon completion of a follow-up article.*

THE remainder of this article is organized as follows. Section I discusses prior research on lease capitalization. Section II includes the development of hypotheses. A description of the methodology is in section III. Section IV analyzes the results. A summary and comments are given in the last section.

I. Background and Prior Research

Potential Economic Effects of Lease Capitalization

SFAS No. 13 (FASB 1976) tightened the conditions under which a lease can be classified as an "operating lease." Furthermore, retroactive capitalization was required for off-balance sheet leases meeting the criteria specified in SFAS No. 13. Capitalizing off-balance sheet leases increases assets and liabilities and typically reduces reported income and retained earnings.[2] These changes increase the likelihood of violating debt covenant restrictions if debt agreements use rolling GAAP.

In the due process hearings prior to the issuance of SFAS No. 13, many lessees claimed that capitalizing off-balance sheet leases would cause violations of debt covenants (Abdel-khalik 1981; Nakayama et al. 1981). In fact, El-Gazzar et al. (1986) found that highly-leveraged firms are more likely to use off-balance sheet leasing. However, the results of studies examining the economic effects of lease accounting and disclosure are inconsistent (e.g., Abdel-khalik 1981; Abdel-Khalik et al. 1978; Ro 1978).

Previous Research of Technical Default Costs

Previous research used financial-statement-based surrogates (mainly debt-to-equity ratio) to examine the extent to which mandated accounting changes affect debt covenants with mixed findings (see for example, Collins et al. 1981; Espahbodi et al. 1991; Leftwich 1981; Lys 1984; Salatka 1989). Notwithstanding the findings of individual studies, Christie (1990) shows that aggregating the results of prior research examining the debt covenant hypothesis produces an overall significant explanatory power.

The inconsistency and/or inconclusiveness of the results across studies suggest that the debt-to-equity ratio may not be an adequate proxy for the closeness to debt covenant violations. Little research evidence has been based on actual debt contracts in examining the economic effects of mandatory accounting changes. For example, Frost and Bernard (1989) use the actual debt contracts of a small sample of oil companies to examine the effects of the SEC rule that mandated write-offs of exploration costs in the first quarter of 1986. They conclude that the write-offs caused significant tightness and even some violations in covenant constraints. However, they show no significant market reaction to the adverse effects on financial statements.

---

[2] Footnote disclosures of most lessees in 1976 financial statements, the year before the effective date of SFAS No. 13, state that capitalizing off-balance sheet leases would cause a decrease in net income and/or retained earnings.

Healy and Palepu (1990) use Moody's summaries about dividend restrictions to examine whether managers make income-increasing accounting changes to circumvent near violation of a dividend constraint. They conclude that, except for pension accounting, managers did not use income-increasing accounting changes to avoid dividend cuts.

II. Development of Hypotheses

A typical debt agreement states the types of restrictions imposed on the borrower during the loan (or bond) term and defines the accounting methods used to measure these restrictions. Covenants that employ financial ratios often restrict the payment of dividends and the issuance of additional debt. In addition, some specify minimum working capital, current ratio, and interest-coverage requirements (Kalay 1982; Smith and Warner 1979).

Retroactive capitalization of leases, formerly classified as operating, alters financial ratios used for calculating covenant restrictions. In this study, the percentage increase in covenant tightness from capitalizing off-balance sheet leases is measured according to the following equation:

$$PITR_{ij}=[(DDVB_{ij}-DDVA_{ij})/(DDVB_{ij})], \tag{1}$$

where:

$PITR_{ij}$ = percentage increase in the tightness of restriction i ($i=1,2,\ldots,k$) for firm j ($j=1,2,\ldots,N$). This also represents the reduction in the slack (distance) to default value of restriction i for firm j,

$DDVB_{ij}$ = distance from the default value of restriction i for firm j *before* capitalizing off-balance sheet leases, and

$DDVA_{ij}$ = distance from the default value of restriction i for firm j *after* capitalizing off-balance sheet leases.

Technical Default Hypotheses

Lessees with significant increases in the tightness of covenant constraints might incur additional costs as a result of operating under tighter contractual conditions or actual technical default. If lenders do not waive the violations, management could take actions to alleviate these effects. These actions could include issuance of additional equity capital, redemption of outstanding debt, or renegotiation of debt contracts. All of these actions are costly in that they alter the mix of investing and financing policies that is based on other economic analysis.

Assuming a nonzero cost of renegotiation or default, affected lessees would incur the cost of either renegotiating terms of the offending debt or operating under tighter contractual conditions. Thus, lessees experiencing a significant increase in the tightness of covenant constraints are negatively affected by SFAS No. 13. Accordingly, the following hypotheses (in the alternative form) are tested:

H1: There is a significant market reaction for affected lessees to the different events in the regulatory process that led to the issuance of SFAS No. 13.

H2: The impact of mandating SFAS No. 13 on securities of affected lessees is negatively related to the percentage increase in the tightness of debt covenant restrictions ($PITR_{ij}$).

Differential Negotiation/Default Cost Hypothesis

The Trust and Indenture Act of 1939 requires that a trustee be appointed for each public debt issue. The trustee is limited in accepting any revision or modification in the debt agreement, which makes it difficult for the manager to renegotiate a defaulted public debt contract, especially if it is widely held.

Based on the above rationale, Leftwich (1981) and Collins et al. (1981) predict higher renegotiaion costs for public debt compared to private debt. However, their results do not support the prediction, which may reflect the inadequacy of the surrogates employed or that the structure of public debt covenants is less dependent on accounting variables. The validity of these inferences can be clarified by analyzing actual debt agreements using direct measures of closeness to default developed for a sample of actual public and private debt contracts. The public-private dichotomy suggests the following hypothesis (in the alternative form):

H3: The correlation between changes in security prices and the increase in tightness of covenant restrictions is stronger for public than for private debt contracts.

III. Sample and Methodology

Test Events

The events that preceded the issuance of SFAS No. 13 by the *Financial Accounting Standards Board* (FASB) and the expected market reaction to each are summarized in table 1. Events 1, 2, 3, 6, and 7 are considered "bad news" because they provided information indicating eventual capitalization of off-balance sheet leases. Event 4 is considered uncertain because the news indicated ambiguity about whether the FASB would reconsider its prior pro-capitalization. The content of event 5 is mixed. Since *The Wall Street Journal* (WSJ) article reporting this event states that the FASB is back to the drawing board, it may be interpreted as "good news" in that firms would be allowed more opportunity to lobby against lease capitalization. On the other hand, the article states that members of the FASB insist on retroactive application of the forthcoming rule which would be considered "bad news." The effect of this event, therefore, is not clear and tests of events 4 and 5 are nondirectional tests.

Sample

Companies included in the AICPA's *Accounting Trends and Techniques* (ATT) for fiscal year 1976 represent the base sample of this study. ATT provides a breakdown of 600 companies by lease-accounting method. Of these, 300 companies used operating lease accounting, while 41 used capital lease accounting for all of their leases. These

Table 1

Summary of the Events Leading up to the Issuance of SFAS No. 13

Event Number	*Content*	*Expected Market Reaction*
1	The Discussion Memorandum July 10, 1974. The FASB issued a discussion memorandum on accounting for leases and set public hearings for November 18, 1974 in New York.	Negative (bad news)
2	The Public Hearings Effects November 25, 1974. Opponents of lease-accounting changes fail to convince the FASB of the dangers of tightening lease accounting.	Negative (bad news)
3	First Exposure Draft August 29, 1975. The FASB exposure draft proposed new rules in the controversial area of lease reporting.	Negative (bad news)
4	Issuance of Statement Delayed November 26, 1975. Delay of adoption of final rules until next year announced.	Uncertain (good news/ bad news)
5	Modifications of the Exposure Draft June 3, 1976. FASB is back to the drawing boards after failing to reach consensus on proposed elimination of retroactive application of new standard. Unspecified changes to draft reported.	Uncertain (good news/ bad news)
6	Second Exposure Draft Issued July 28, 1976. New rule will apply to existing leases as well as new leases.	Negative (bad news)
7	Final Statement Issued December 3, 1976. Major changes implemented in leasing.	Negative (bad news)

Note: Events are identified by *The Wall Street Journal* article since the FASB meetings were confidential in the earlier years and only the minutes of the meetings were released to the press on later dates.

two groups had chosen to use exclusively one method of accounting for leases before SFAS No. 13. Other ATT listed firms either used a mix of accounting methods for leases or did not provide adequate disclosure.

Examination of the annual reports of the 300 noncapitalizing firms revealed that 113 subsequently restated financial statements.[3] Another 125 firms did not restate financial statements due to immaterial effects of the change in the lease accounting method. The remaining 62 firms ceased to report as a separate entity in the years following 1976.[4] The 41 capitalizers and the 113 restatement companies were further reduced to 23 and 60 by the selection limits reported in table 2.[5]

---

[3] Originally, the FASB granted a long transition period for retroactive application of SFAS No. 13. However, the SEC shortened this period to no later than 1978. The majority of sample firms had adopted SFAS No. 13 retroactively in 1978.

[4] This category represents companies that were delisted for acquisitions, firms that made quasi-reorganization and changed names, and firms that went bankrupt. This number is not unusual since the ATT shows that 83 companies were delisted/substituted mainly because of mergers and acquisitions over the same period investigated in this study, fiscal years 1977–1979.

[5] An analysis of sample characteristics, that is not reported here, shows that the sample firms are well-dispersed over the two-digit industry classification.

Table 2
Final Sample Selection Process

	Control Sample: All Leases Capitalized	*Experimental Sample: Operating Leases/ Restated Retroactively*
Initial sample	41	113
Minus:		
Companies with dividends or earnings announcements	3	14
Companies not available on CRSP daily stock return	14	22
Companies for which no debt contracts found	NA	14
Companies with reorganization or bankruptcy activities	1	3
Final sample	23	60

NA=Not Applicable

Debt Contracts: Cross-Sectional Test Sample

Debt contracts were needed only for the experimental group on which cross-sectional tests were performed. Debt contracts that existed before November 30, 1976 were obtained from statements of debt registrations, filed prospectus reports, and 10-K filings of 1980 and 1981 that contain detailed contracts for the 1976 debt.[6] This study uses the most restrictive covenants of known contracts.[7]

To the extent that test events are on different dates, the number of contracts (sample size) for each event is different, reflecting issuance of new contracts or settlement of old ones. Both the market and the cross-sectional tests are conducted for the cumulative effect of all events using all available debt contracts.

Coding and Analysis of Debt Agreements

Coding of debt contracts is summarized in table 3. For the sample including all debt agreements, the most binding contract is used. Accounting-based restrictions contain 47 restrictions on issuing additional debt and 42 on payment of dividends. Twelve contracts require maintenance of certain levels of working capital, while seven contracts require maintenance of a minimum current ratio and four contracts have restrictions on fixed-charges coverage. The coding of public and private debt contracts reveals few differences in the use of accounting-based covenants.

[6] The Form 10-Ks for 1980 and 1981 represent a special source of debt agreements because many firms attached copies of debt contracts as exhibits to financial statements in compliance with the SEC integrated disclosure act of 1980.

[7] Lack of the most restrictive debt contracts produces bias against finding significant correlation between the tightness of covenant restrictions and the market returns associated with SFAS No. 13. Since the results show significant correlations, no further effort is necessary to identify those contracts.

Table 3
Coding of Debt Agreements of Affected Lessees

	Type of Restriction (Frequency/Percentage %)			
	Negative			*Affirmative*
Sample	*Total*	*ADR*	*DPR*	*WC/CR*
All Debt Agreements (one contract per company)				
Contracts with no accounting-based restrictions	4 (7%)			
Contracts with accounting-based restrictions but adjustment for off-balance sheet leases	4 (7%)	3 (6%)	2 (5%)	
Contracts with accounting-based restrictions and no adjustment for off-balance sheet leases	52 (86%)	44 (94%)	40 (95%)	19 (100%)
Total	60 (100%)	47 (100%)	42 (100%)	19 (100%)
Private Debt Agreements Only				
Contracts with no accounting-based restrictions	2 (6%)			
Contracts with accounting-based restrictions but adjustment for off-balance sheet leases	3 (10%)	2 (8%)	2 (9%)	
Contracts with accounting-based restrictions and no adjustment for off-balance sheet leases	25 (84%)	23 (92%)	21 (91%)	12 (100%)
Total	30 (100%)	25 (100%)	23 (100%)	12 (100%)
Public Debt Agreements Only				
Contracts with no accounting-based restrictions	3 (8%)			
Contracts with accounting-based restrictions but adjustment for off-balance sheet leases	2 (5%)	2 (7%)	1 (4%)	
Contracts with accounting-based restrictions and no adjustment for off-balance sheet leases	32 (87%)	26 (93%)	24 (96%)	7 (100%)
Total	37 (100%)	28 (100%)	25 (100%)	7 (100%)

ADR = additional debt restrictions;
DPR = dividend payment restrictions; and
WC/CR = restrictions requiring certain level of working capital and/or current ratio. These are grouped in one category because they measure the same attribute, short-term liquidity, and because the number of contracts requiring these restrictions is very small.

Table 4
Calculated Percentage Increase in the Tightness of Covenant Restrictions

	Cross-Sectional Distribution Statistics					
Variable	*Mean*	*S.D.*	*25%*	*50%*	*75%*	*100%*
All Debt Agreements						
PITR1	43.1	29.9	25.9	45.4	68.6	92.4
PITR2	18.3	18.9	8.2	20.3	37.8	77.2
PITR3	7.9	2.3	0.0	0.0	11.5	65.0
Private Debt Agreements						
PITR1	54.1	30.0	31.0	61.0	76.1	92.4
PITR2	21.5	22.0	4.4	19.4	42.8	77.2
PITR3	NM	NM	NM	NM	NM	NM
Public Debt Agreements						
PITR1	37.2	25.4	0.0	34.7	37.3	79.5
PITR2	15.6	16.0	9.8	13.2	18.8	67.3
PITR3	NM	NM	NM	NM	NM	NM
Differences between Means (private vs. public)						
PITR1	16.9*					
	(1.69)					
PITR2	5.9**					
	(1.59)					

PITR1 = percentage increase in the tightness of restrictions on issuance of additional debt,
PITR2 = percentage increase in the tightness of restrictions on payment of dividends,
PITR3 = percentage increase in the tightness of restrictions requiring maintenance of certain levels of working capital and/or current ratio, and
NM = not meaningful since the frequencies are very low when partitioned into private and public (see table 3).

* Significantly different from zero at 0.05 or less.
** Significantly different from zero at 0.10 or less.

Calculated Effects of SFAS No. 13 on Covenant Tightness

Table 4 presents the percentage increase in tightness of covenant restrictions that would have occurred if off-balance sheet leases were capitalized (see the appendix). These calculations are made using data from footnotes on lease rent commitments as reported in 1976 financial statements under the mandates of ASR No. 147 (SEC 1973).

For the category of all debt agreements, the mean (median) increase in tightness of restrictions on additional debt is 43.1 (45.4) percent and is 18.3 (20.3) percent on payments of dividends. The mean increase in working capital and/or current ratio is 7.9 percent. The distribution of the working capital variable is significantly skewed to the right, apparently due to the infrequent use of this restriction in sample contracts.

In addition, covenant constraints of private debt seem tighter than public debt. The mean (median) increase in tightness of private debt covenants on issuing new debt is 54.1 (61.0) percent, and is 21.5 (19.4) percent on payments of dividends. The corresponding increases for public debt are 37.2 (34.7 percent for issuing new debt and 15.6 (13.2) percent for dividends.[8]

Return Generating Model

The market reaction to the events of SFAS No. 13 is measured by the market model with a modification to include dummy variables to extract the effect of the event on the firm's realized returns (see Hughes and Ricks 1984; Schipper and Thompson 1983). Each dummy variable takes the value of *one* on the date of the *k*th event and zero otherwise.[9] This model takes the following form:

$$(R_{jt}-R_{0t})=A_j+B_j(R_{mt}-R_{0t})+\sum_k C_{jk}(P_{kt})+U_{jt}, \tag{2}$$

$$t=1,2,\ldots,T;\quad j=1,2,\ldots,N;\quad k=1,2,\ldots,k,$$

where:

R_{jt} = return on stock of firm j during day t,
R_{mt} = return on the value weighted portfolio during day t as compiled by CRSP,
R_{0t} = return on short-term U.S. Treasury bills on day t,
A_j, B_j = intercept and slope parameter of firm j, respectively,
C_{jk} = returns effect of the kth event on the jth firm stock, $k=1,2,\ldots,7$,
P_{kt} = a dummy variable taking the value of one in kth event period $(-1,0)$ and zero otherwise, where day (0) denotes the publication of *The Wall Street Journal* article commenting on the event. Thus, P_{kt} is a vector consisting of zeros in all but the two-day period of the kth event, and
U_{jt} = a disturbance term of firm j on day t, and is assumed to be $NID(0,\sigma_j)$.

Equation (2) is also estimated for equally weighted portfolios of sample firms, which allows for efficient use of available data in estimating the model parameters A_j and B_j. The t-values of the regression coefficients provide a direct test of the significance of estimated abnormal returns (C_k) for the events under study.

The Cross-Sectional Model

Changes in the tightness of covenant restrictions are related to changes in security prices through the estimated event parameters, C_{jk}, as follows:

$$\begin{aligned}C_{jk}=&A_0+A_1(PITR1)_{jk}+A_2(PITR2)_{jk}+A_3(PITR3)_{jk}+A_4(PUB^*PITR1)_{jk}\\&+A_5(PUB^*PITR2)_{jk}+A_6(PUB^*PITR3)_{jk}+A_7(SUBORD)_{jk}+e_{jk},\end{aligned} \tag{3}$$

---

[8] Percentage increase in the tightness of a given restriction is affected by the magnitude of off-balance sheet leases and by the default value of the restriction. Therefore, it is possible that observed differences between public and private debt are attributable to differences in magnitude of off-balance sheet leases. To investigate that, the ratio of noncapitalized leases to outstanding debt, and the ratio of the amount of decline in retained earnings to retained earnings before capitalization were calculated. Because the mean and median of these ratios are not significantly different for public and private debt, it was concluded that the observed differences in increase to default are the result of the stated default value of private versus public contracts.

[9] Results of Seemingly Unrelated Regressions (not reported) show no significant differences from the results of the market reaction test used in this study.

where:

C_{jk} = the abnormal return of firm j for the kth event, calculated with equation (2),

$PITR1_{jk}$ = percentage increase in the tightness of restrictions on additional debt of firm j,

$PITR2_{jk}$ = percentage increase in the tightness of restrictions on payment of dividends of firm j,

$PITR3_{jk}$ = percentage increase in the tightness of restrictions on working capital, current ratio, and/or interest coverage of firm j,

PUB_j = a dummy variable representing public debt. It takes a value of one if the closeness measure is from a public contract, and a zero value otherwise,

$SUBORD_j$ = a dummy variable representing subordination of the debt. It takes a value of one for subordinated debt and zero for secured, sinking fund, and senior debt, and

e_{jk} = a stochastic disturbance term.

IV. Results

Average Abnormal Returns

Table 5 presents regression estimates of equation (2) for equally weighted portfolios. For the test sample, the event parameters were significantly negative for event 2 (FASB public hearings on the subject) and event 5 (the WSJ article stating that the FASB is back to the drawing board). The test sample has an average abnormal return of -0.47 percent for event 2 ($P<0.10$) and -0.68 percent for event 5 ($P<0.05$).

The results for event 5 indicate that market participants interpreted the WSJ article as "bad news." Given that the *ex ante* classification of this event is uncertain, it is also possible that the observed negative market reaction was driven by other activities. However, further examination of the WSJ on the dates of the events did not reveal any evidence to substantiate this concern. In relation to the control sample, the test sample has an average cumulative abnormal return of -0.163 percent, which is significantly different from zero at $p<0.06$.

The Cross-Sectional Results

Results on Technical Default Hypotheses. Table 6 presents regression estimates of the cross-sectional tests for events 2 and 5, as well as for the cumulative market effect of all events. The coefficient of the increase in tightness of restrictions on issuance of additional debt (A_1) maintains a negative sign and is significant in all tests. The coefficient of the increase in the tightness of the dividend payment restriction (A_2) has a negative sign in two tests but is significant only for event 5.[10] The effect of the increase in the tightness of restrictions on working capital (A_3) maintains a negative sign but is insignificant in all tests.[11]

---

[10] One should be careful when interpreting the coefficients of the cross-sectional regressions in that they are not additive. The events and variables are mutually exclusive since defaults in one event are eliminated in the following events. Further, although the two default measures, additional debt and dividend, may be significant in the same event, one makes the other redundant; i.e., violating one restriction can lead to technical default.

[11] Regression results using the debt-to-equity ratio as the surrogate for debt covenant tightness, did not reveal a significant relationship to the abnormal return of sample firms.

Table 5
Regression Estimates of Market Reaction to Events Preceding SFAS No. 13

(*t*-values in parentheses)

$$\text{Model: } 1/N\sum_{j}(R_{jt}-R_{0t})=\bar{A}+\bar{B}(R_{mt}-R_{0t})+\sum_{k}\bar{C}_k(P_{kt})+U_t$$

Event Number	*Noncapitalizers (Test Sample)*	*Capitalizers (Control Sample)*
1	0.00061 (0.232)	−0.00103 (−0.092)
2	−0.00472** (−1.790)	0.00267 (0.238)
3	0.00018 (0.069)	0.00161 (0.143)
4	0.00049 (0.187)	0.00141 (0.126)
5	−0.00675* (−2.558)	−0.00007 (−0.006)
6	0.00105 (0.397)	−0.00181 (−0.161)
7	−0.00037 (−0.143)	−0.00254 (−0.228)
Cumulative Average Abnormal Return	−0.00135*** (−1.374)	0.00028 (0.077)

* Significant at 0.05 level or less, two-tailed test.
** Significant at 0.05 level or less, one-tailed test.
*** Significant at 0.10 level or less, one-tailed test.

Results on the Differential Negotiation/Default Costs. Table 6 also presents test results for the effect of the public-private dichotomy on abnormal returns, which relies on estimates of interaction terms. The results are not consistent with the hypothesis; the coefficients of the public debt variable (A_4, A_5, and A_6) exhibit mixed signs and are insignificant in all tests, suggesting that public and private debt may not differ in terms of the effects of the increase in the closeness of covenants to their constraints. These results may be due to other factors. For example, El-Gazzar and Pastena (1990) found that the lender's industry and syndication of debt influence how accounting rules are tailored in private debt contracts. Thus, other nonaccounting provisions in debt covenants seem to affect the scope and tightness of accounting-based covenant restrictions.

Interaction with Nonaccounting-Based Provisions

To investigate the inconsistency of the results of the public-private dichotomy with expectations, nonaccounting-based provisions in each contract are analyzed. Results are presented in table 7.

The results show that public debt has a higher percentage of contracts using sinking fund or security provisions. In contrast, private debt has a higher percentage of contracts characterized as subordinated debt or revolving credit agreements. These results

Table 6

Regression Estimates of the Cross-Sectional Relationship Between Changes in Security Prices and Changes in the Tightness of Covenant Restrictions
(t-values in parentheses)

Model: $C_{jk}=A_0+A_1(PITR1)_j+A_2(PITR2)_j+A_3(PITR3)_j+A_4(PUB^*PITR1)_j$
$+A_5(PUB^*PITR2)_j+A_6(PUB^*PITR3)_j+A_7(SUBORD)_j+e_j$

Coefficients (Expected Sign)	Event 2		Event 5		Cumulative Effect (All Events)	
	Run 1	Run 2	Run 1	Run 2	Run 1	Run 2
A_0	0.0079***	0.0093*	−0.0013	0.0005	0.0151***	0.0203*
(?)	(1.51)	(2.02)	(−0.37)	(0.15)	(1.59)	(2.36)
A_1	−0.0462*	−0.0398*	−0.0162**	−0.0091**	−0.0740*	−0.0506*
(−)	(−3.55)	(−5.14)	(−1.86)	(−1.79)	(−3.15)	(−3.67)
A_2	0.0507	0.0138	0.0044	−0.0147*	0.0676	−0.0178
(−)	(1.28)	(1.27)	(0.17)	(−2.02)	(0.94)	(−0.91)
A_3	−0.1075	−0.0110	−0.0061	−0.0074	−0.0048	−0.0072
(−)	(−0.90)	(−0.81)	(−0.47)	(−0.77)	(−0.14)	(−0.28)
A_4	0.0102	NU	0.0115	NU	0.0347	NU
(−)	(0.65)		(1.11)		(1.28)	
A_5	−0.0389	NU	−0.0194	NU	−0.0896	NU
(−)	(−0.95)		(−0.71)		(−1.20)	
A_6	0.0380	NU	0.0156	NU	0.0456	NU
(−)	(1.09)		(0.78)		(0.77)	
A_7	−0.0004	−0.0001	−0.0038	−0.0047***	−0.0130***	−0.0158**
(−)	(−0.07)	(−0.04)	(−1.10)	(−1.53)	(−1.39)	(−1.84)
Adjusted R-squared	.32	.31	.14	.13	.23	.22

C_{jk}=the market reaction of firm j to event k calculated by using equation (2);
PITR1=percentage increase in the tightness of restrictions on issuance of additional debt;
PITR2=percentage increase in the tightness of restrictions on payment of cash dividend;
PITR3=percentage increase in the tightness of restrictions on working capital and/or current ratio;
PUB=a dummy variable with a value of 1 for public debt contracts and a value of 0 for private contracts;
SUBORD=a dummy variable with a value of 1 for subordinated debt and 0 for secured, senior, or sinking fund debt; and
NU=not used.

* Significant at 0.01 or less, one-tailed test.
** Significant at 0.05 or less, one-tailed test.
*** Significant at 0.10 or less, one-tailed test.

seem to provide some support for the relevance of the coefficient of the subordinated variable presented in table 6; the coefficient of this variable (A_7) maintains the expected sign in all tests and is significant in two of the three.

V. Conclusions and Recommendations

Previous research, which assumes that the probability of technical default varies with leverage ratios, produced mixed results. In contrast, this article uses actual debt

Table 7
Summary of Major Nonaccounting Provisions in Debt Contracts

Provisions	*Full Sample* (n=60)	*Public* (n=37)	*Private* (n=30)
Senior debt	28 (46%)	19 (51%)	9 (30%)
Sinking fund	31 (49%)	26 (70%)	5 (17%)
Security	20 (33%)	16 (43%)	4 (13%)
Subordinated debt	29 (48%)	12 (32%)	18 (60%)
Convertible debt	14 (22%)	11 (29%)	3 (10%)

Note: One agreement per company is included. When the company has both public and private contracts, the most restrictive one is used. Thus, 60 of the 67 debt contracts are included in the full sample.

contracts to develop a more precise measure of the impact of SFAS No. 13 on the tightness of debt covenant restrictions, and to examine its effects on abnormal returns.

The results are consistent with the hypothesis that capitalizing off-balance sheet leases increased the probability of technical default. The percentage increase in the tightness of debt covenant restrictions is correlated with changes in security prices that accompanied the regulatory events of SFAS No. 13. In addition, the analysis of debt contracts shows some findings that have implications for the differentiation between private and public debt. Private debt agreements have tighter restrictive financial covenants, while public debt contracts utilize more nonaccounting provisions such as security and/or a sinking fund and seek seniority status of the debt.

Appendix

Calculations of the Percentage Increase in the Tightness of Covenant Constraints from Compliance with SFAS No. 13

The calculation of the percentage increase in the tightness of covenant constraints is made by using equation (1) as follows:

$$PITR_{ij}=[(DDVB_{ij}-DDVA_{ij})/(DDVB_{ij})]$$

where:

$PITR_{ij}$=percentage increase in the tightness of restriction i ($i=1,2,\ldots,k$) for firm j ($j=1,2,\ldots,N$). This also represents the reduction in the slack (distance) to default value of restriction i for firm j.

$DDVB_{ij}$=distance from the default value of restriction i for firm j *before* capitalizing off-balance sheet leases.
=Restriction j Default Value−Covenant Based Measure of Restriction j *Before* Capitalizing Off-Balance Sheet Leases.

$DDVA_{ij}$=distance from the default value of restriction i for firm j *after* capitalizing off-balance sheet leases.
=Restriction j Default Value−Covenant Based Measure of Restriction j *After* Capitalizing Off-Balance Sheet Leases.

To calculate $DDVB_{ij}$ and $DDVA_{ij}$, one needs the following items: the definition and limits of restriction i from the debt contract of firm j, the financial data from financial statements, and the note disclosure on lease obligations (present value of noncapitalized leases and the impact on net income had these leases been capitalized).

Assume the following example. XYZ Company has a debt contract stating the following restriction on issuance of additional debt: "the ratio of long-term debt to tangible net worth should not exceed 1.5." The contract defines long-term debt as all liabilities maturing beyond one year, and tangible net worth as total shareholders' equity less intangibles.

The following disclosure on the lease obligations was stated in the 1976 annual report of the company (under the SEC's disclosure mandate, ASR No. 147): the company leases some properties and equipment that are accounted for under operating lease, and the present value of noncapitalized leases and the impact on net income had these leases been capitalized are $20 and $0.5 millions, respectively.

The balance sheet of XYZ on 12/31/1976 is as follows (in millions):

Assets		*Liabilities and Equity*	
Current assets	50	Current liabilities	35
Plant assets	150	Long term debt	85
Intangibles	15	Shareholders' equity	95
Total assets	215	Total liabilities and equity	215

Calculations

1. The ratio of debt to net worth *before* capitalization $=(85)/(95-15)=1.06$.
2. $DDVB_{ij}=1.5-1.06=0.44$.
3. The ratio of debt to net worth after capitalizing off-balance sheet leases $=(85+20)/(95-15-0.5)=105/79.5=1.32$.
4. $DDVA_{ij}=1.5-1.32=0.18$.

Accordingly, $PITR_{ij}=(0.44-0.18)/0.44=0.59$. Therefore, retroactive capitalization of off-balance sheet leases would have reduced the slack of the restriction on issuance of additional debt by 59 percent.

References

Abdel-khalik, A. R. 1981. *The Economic Effects on Lessees of FASB Statement No. 13, Accounting for Leases*. Research report, Stamford, CT: FASB.

——, R. Thompson, and R. Taylor. 1978. The impact of reporting leases off the balance sheet on bond risk premiums. In *Economic Consequences of Financial Accounting Standards: Selected Papers*, Stamford, CT: FASB, 101–58.

Christie, A. 1990. Aggregation of test statistics: An evaluation of the evidence on contracting and size hypotheses. *Journal of Accounting & Economics* 12 (January): 15–36.

Collins, D. W., M. S. Rozeff, and D. Dhaliwal. 1981. The economic determinants of market reaction to proposed mandatory accounting changes in the oil and gas industry: A cross-sectional analysis. *Journal of Accounting & Economics* 3 (March): 37–71.

El-Gazzar, S. M., S. Lilien, and V. Pastena. 1986. Accounting for leases by lessees. *Journal of Accounting & Economics* 8 (October): 217–37.

——, and V. Pastena. 1990. Negotiated accounting rules in private financial contracting. *Journal of Accounting & Economics* 12 (March): 381–96.

Espahbodi, H., E. Strock, and H. Tehranian. 1991. Impact on equity prices of pronouncements related to nonpension postretirement benefits. *Journal of Accounting & Economics* 14 (December): 323–46.

Financial Accounting Standards Board. 1976. *Statement of Financial Accounting Standards No. 13: Accounting for Leases*. Stamford, CT: FASB.

Frost, C., and V. Bernard. 1989. The role of debt covenants in assessing the economic consequences of limiting capitalization of exploration costs. *The Accounting Review* 64 (October): 788–808.

Healy P., and K. Palepu. 1990. Effectiveness of accounting-based dividend covenants. *Journal of Accounting & Economics* 12 (January): 97–124.

Hughes, J., and W. Ricks. 1984. Accounting for retail land sales: Analysis of a mandated change. *Journal of Accounting & Economics* 6 (August): 101–32.

Kalay, A. 1982. Stockholder-bondholder conflicts and dividend constraints. *Journal of Financial Economics* 10 (July): 211–33.

Leftwich, R. 1981. Evidence of the impact of mandatory changes in accounting principles on corporate loan agreements. *Journal of Accounting & Economics* 3 (March): 3–36.

Lys, T. 1984. Mandated accounting changes and debt covenants: The case of oil and gas accounting. *Journal of Accounting & Economics* 6 (April): 39–65.

Nakayama, M., S. Lilien, and M. Benis. 1981. Due process and FASB No. 13. *Management Accounting* 62 (April): 49–53.

Ro, B. T. 1978. The disclosure of capitalized lease information and stock prices. *Journal of Accounting Research* 16 (Autumn): 315–40.

Salatka, W. 1989. The impact of SFAS No. 8 on equity prices of early and late adopting firms: An events study and cross-sectional analysis. *Journal of Accounting & Economics* 11 (February): 35–69.

Schipper, K., and R. Thompson. 1983. The impact of merger-related regulations on the shareholders of acquiring firms. *Journal of Accounting Research* 21 (Spring): 184–221.

Securities and Exchange Commission. 1973. ASR No. 147, *Notice of Adoption of Amendments to Regulation S-X Requiring Improved Disclosure of Leases*. Chicago: Commerce Clearing House.

Smith, C., and J. Warner. 1979. On financial contracting: An analysis of bond covenants. *Journal of Financial Economics* 7 (June): 117–61.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 273–288

Debt Contracts and FAS No. 19: A Test of the Debt Covenant Hypothesis

Mary Beth Mohrman
University of Missouri–St. Louis

SYNOPSIS AND INTRODUCTION: The objective of this study is to evaluate the debt covenant hypothesis by using the details of lending agreements at the time of the FAS No. 19 Exposure Draft. If adopted, FAS No. 19 would have eliminated the use of the full-cost method for firms involved in oil and gas exploration.

Several studies found a significant difference between the abnormal returns of full-cost and successful-efforts firms in response to the release of the FAS No. 19 Exposure Draft (Collins and Dent 1979; Lev 1979; Lys 1984), and this has been hypothesized to result from possible effects on accounting-based covenants in debt contracts. However, empirical tests were indirect in that variables of financial leverage were used as proxies for debt covenant effects.

According to the debt covenant hypothesis, firms choose accounting methods to maximize slack in debt covenant constraints. Mandatory changes in accounting methods might bring some firms closer to violating their debt covenants. Borrowing firms, however, can anticipate the possibility of such regulatory change and protect themselves, for example, by completely specifying the accounting methods that will be used to determine covenant compliance, regardless of subsequent changes to generally accepted accounting principles. When outstanding debt contracts are so structured, the probability of default on accounting-based covenants could be unaffected by GAAP changes. Financial leverage variables fail to distinguish between firms with and without such protection.

Using exhibits to SEC filings, I examined 83 debt contracts in effect in

This article is based on my dissertation research at the John M. Olin School of Business, Washington University, St. Louis, MO. I am indebted to my dissertation committee members, Nick Dopuch, Mort Pincus, and Carol Frost, for their guidance and suggestions. Greg Waymire and Grace Pownall provided helpful comments in earlier stages of this research. I also thank Silvia Madeo, two anonymous reviewers, and an associate editor for their comments. Financial assistance from the Deloitte Haskins and Sells Foundation and the John M. Olin School of Business is gratefully acknowledged.

Submitted August 1991.
Accepted December 1992.

1977 for each firm in a sample of 35 full-cost firms. I identified 13 firms that had at least one covenant that would have been affected by FAS No. 19. Portfolio tests indicated that these 13 firms drive the difference between the returns of full-cost and successful-efforts firms in the full sample. An alternate set of portfolios based on leverage as a proxy for debt covenant effects showed no relation to those based on actual debt contracts and also failed to explain the stock price response to the Exposure Draft release.

Key Words: *Debt covenants, Regulation, Capital markets.*

Data Availability: *A list of sample firms and debt contracts is available from the author. Stock price and financial statement data are available from public sources.*

THE remainder of this article is organized as follows. In section I, I discuss FAS No. 19 and describe related prior studies. In section II, I describe the sample selection and data collection procedures. In section III, the analysis of the debt contracts and subsequent partitioning of the full-cost firms is presented; in section IV, the stock price and leverage tests, based on two groups of full-cost firms are presented. A summary and conclusions appear in section V.

I. FAS No. 19 and Prior Studies

In 1977, the Financial Accounting Standards Board issued the *Exposure Draft for Financial Accounting Standard No. 19*: Financial Accounting and Reporting by Oil and Gas Producing Companies. The proposed standard would have eliminated the full-cost method and required the use of a form of the successful-efforts method for all oil and gas firms.

The effect of the proposed regulation on firm values was the subject of several studies. Studies using weekly abnormal returns produced conflicting results. Collins and Dent (1979) found a significant difference between the weekly abnormal returns of full-cost and successful-efforts firms around the event period, but Dyckman and Smith (1979) did not find that difference to be significant. However, studies using daily abnormal returns (Lev 1979; Lys 1984) confirmed Collins and Dent's result.

Several studies interpreted this reaction as a debt covenant effect without testing (Lev 1979) or with indirect tests that used leverage ratios and indicator variables (Collins et al. 1981; Lys 1984). Lys included a debt/equity ratio and a proxy for the variance of each firm's cash flows, and the coefficient estimates for both were significant. Lys (1984, 61) concluded that, "debt covenants are important and ... these contracts can have an impact on the wealth of shareholders when accounting changes occur."

These studies and others have ignored the details of debt contracts. Foster (1980) examined footnote disclosures for full-cost firms and identified several firms that were included in previous studies even though they had reported that FAS No. 19 would have no effect on their debt covenant compliance. If the debt covenant explanation is valid, firms that did experience debt covenant effects would have exhibited more significantly negative abnormal returns than those that did not. Unfortunately, only a small group of full-cost firms included information about debt covenant compliance in their financial statement footnotes. Identifying the full-cost firms that would have experienced debt covenant effects requires analysis of actual debt contracts.

In this study, examination of actual debt contracts provided the basis for partitioning a sample of full-cost firms into those that would have been affected by FAS No. 19 and those that would not. I test the hypothesis that firms with such effects experienced more negative abnormal returns at the time of the FAS No. 19 Exposure Draft. I also partitioned the firms on the basis of leverage and tested the hypothesis that leverage is an adequate substitute for examination of contracts to determine debt covenant effects.

II. Sample Selection and Data Collection

Two samples were selected: a sample of successful-efforts firms with 1977 stock price data available and a sample of full-cost firms with the same stock price data and debt contracts available for analysis. A search of SIC codes in the CRSP data base revealed 129 firms likely to have been involved in oil and gas exploration with daily stock price returns available on the NYSE/AMEX CRSP tapes for the period January 3 to December 30, 1977.[1] Of these firms, the 1977 Form 10-K reports were unavailable for 12, and 13 of the firms became involved in oil and gas exploration after 1977. The 44 firms that used the successful-efforts method in 1977 formed the successful-efforts portfolio.

The remaining 60 firms used the full-cost method in 1977. I examined item 14 in the 10-K reports[2] for fiscal years 1970–1978 for each of these firms, noting any debt contracts listed as exhibits to the Form 10-K or any other SEC filing. The reports contained references to approximately 150 debt contracts for 42 of the 60 firms. I was able to locate 83 debt contracts that were in effect at the time of the FAS No. 19 Exposure Draft for 35 of these firms.[3] These 83 contracts are the basis for the descriptive evidence presented in section III, and the 35 firms form the full-cost portfolio.

III. Analysis of Contracts

The 83 contracts examined included 12 public issue indentures and 71 private debt agreements. The private debt agreements included bank term loan and revolving credit agreements, bond purchase agreements, and promissory notes. Twelve of the 83 contracts—four public issues and eight private placements—contained no covenants using numbers from the borrower's financial statements (i.e., accounting-based covenants).

Of the contracts with accounting-based covenants, 29 contained provisions to protect the terms of the contract in the event of a change in accounting standards. Some specify that covenant compliance will be assessed with all accounting methods in use in a particular set of the firm's financial statements. Others specify GAAP as in effect on the contract date, but allow changes within that set of rules. Still others allow the borrowing firm to choose between GAAP in effect on the contract date and GAAP in effect at the time that compliance is being determined. In addition, some contracts contained an agreement to revise the terms of accounting-based covenants in the event of a change in GAAP. Examples of each of these types of provisions are provided in the appendix. Any of these provisions would allow a firm to comply with covenants by using the full-cost method, even if required to change financial statements to the successful-efforts method. Contracts with any of these provisions are referred to as "frozen GAAP" contracts; others are referred to as "rolling GAAP" contracts.

---

[1] OTC firms were omitted because their 10-K reports were not readily available. Without 10-K reports for 1977, I could not verify that a firm was involved in oil and gas exploration at that time or begin the search for relevant debt contracts. See Lev (1979, 491) for a discussion of additional reasons for excluding OTC firms.

[2] For an extensive description of item 14 and the exhibit and index filing requirements, see Press and Weintrop (1990, 69).

[3] Contracts were located by reading exhibits to Form 10-K reports and other SEC filings. These were available at the SEC Public Reference Room in Washington, D.C., or were purchased on microfiche from Disclosure, Inc. SEC registrants are not required to file debt agreements pertaining to loans having a value of less than 10 percent of the firm's assets. Therefore, many smaller debt issues were not available. In addition, SEC documents filed prior to 1970 were not available at the SEC or from Disclosure, Inc. In some cases, the SEC document was available, but the exhibits had never been photocopied by Disclosure, Inc. The sample used in this study includes all debt contracts that were filed as exhibits and included on the available Disclosure, Inc. microfiche. Of the contracts that were unavailable, seven pertained to two firms that I dropped for lack of contracts. The remaining 45 unavailable contracts pertained to firms that remained in the sample.

The most common accounting-based covenants were those limiting dividends, leverage, working capital, current ratio, net worth, and interest coverage. Working capital and current ratio numbers would not have been affected by FAS No. 19, nor would net worth and leverage constraints if written in terms of tangible assets or tangible net worth because capitalized drilling expenditures were viewed as long term intangible assets. Hence, the analysis concentrated on interest coverage, net worth, leverage, and dividend constraints that would have been affected by FAS No. 19.[4]

For a firm to have experienced changes in debt covenant slack as a result of FAS No. 19, the following necessary conditions must hold: (1) the firm had at least one "rolling GAAP" contract whose covenants used financial statement numbers that would have been affected by a change to the successful-efforts method; (2) FAS No. 19 would have materially affected the firm's net worth and/or retained earnings; (3) when the only affected covenant was a dividend constraint, the borrower was a dividend-paying firm.[5]

I assessed these conditions by examining the 83 contracts including all sections describing rules of construction, definitions, and affirmative and negative covenants (cond. 1); the notes to the financial statements for 1977 for all 35 firms (conds. 1 and 2); and the dividend records of firms with "rolling GAAP" dividend covenants (cond. 3).

The results of this examination are presented in table 1. Analysis of the 83 contracts revealed covenants of 12 firms that would have led to a reduction in slack if FAS No. 19 had become effective. Review of the notes to the 1977 financial statements revealed that one firm was misclassified. Buttes Gas & Oil reported in its financial statement notes that FAS No. 19 would cause loan covenant defaults. On the basis of its report, Buttes was classified as a firm that would have been affected, even though the specific covenants that would have caused default were not identified.[6] Thus, the analysis identified 13 firms that would have experienced debt covenant effects if FAS No. 19 had become effective.

---

[4] Other, less common accounting-based covenants included restrictions on current debt, restrictions on cash flows, and a restriction on investments. None of them would have been affected by FAS No. 19.

[5] It is difficult to assess the impact of a change in dividend covenant slack for a firm that does not pay dividends. If the firm never pays dividends, then a change in the pool of funds available for dividends is irrelevant. If the firm pays dividends in the future, then a reduction in available funds could become relevant. Sensitivity analysis on this point is presented in section IV.

[6] Another firm, Adobe Resources, also reported that FAS No. 19 would have caused a covenant violation. However, the relevant contract was part of my sample, and it contained an agreement to renegotiate the covenants in the event of a GAAP change. In its 1977 10-K report, Adobe reported that the lender had already agreed to waive the restriction.

Table 1
Examination of Contracts

Panel A. Firms with Documented Debt Covenant Slack Changes (FC13):

Firm	*No. of Contracts Examined*	*Remarks*
Apache Corporation	1	Dividend covenant slack reduced.
Belco Petroleum	2	Dividend covenant slack reduced in one contract. Other contract contained no accounting-based covenants.
Buttes Gas & Oil	1	Examined contract contained no accounting-based covenants, but firm reported that it would be in violation of some of its covenants with FAS No. 19.
Coastal States Gas	1	Leverage and dividend slack reduced.
Houston Oil & Minerals	3	One contract used frozen GAAP; dividend slack reduced for remaining two.
Inexco Oil	4	One contract used frozen GAAP; one contained no accounting-based covenants; dividend slack reduced for remaining two.
Mitchell Energy	2	Leverage slack reduced in one contract; dividend slack reduced in other.
Panhandle Eastern	4	One contract contained no accounting-based covenants; dividend and interest coverage slack reduced for remaining three.
Petro-Lewis	2	Net worth slack reduced in one contract; remaining contract used frozen GAAP.
Texas International	7	Four contracts used frozen GAAP; two contained no accounting-based covenants; dividend slack reduced for remaining contract.
Transco Company	1	Already in violation of leverage constraint; dividend slack reduced.
Universal Resources	4	One contract used frozen GAAP; one contained no accounting-based covenants; dividend slack reduced in remaining two.
Williams Companies	2	Dividend slack reduced in one contract.

Continued

Of the 13 firms, seven were affected by the dividend constraint. Dividend constraints are usually negative covenants; that is, once a firm is out of compliance with the terms of a dividend covenant, it is not considered in default unless it pays another dividend. A dividend covenant could not have caused immediate default if FAS No. 19 had become effective. If a dividend constraint became binding, any of these seven firms could have avoided default by reducing or omitting future dividends. However, such a reduction could be costly to stockholders. The value of debt is determined in part by the restrictiveness of its covenants. When debt is issued, stockholders "pay" for the right to receive dividends at a particular level. When a subsequent change of accounting regulations causes a dividend covenant to be more binding than it was intended to be, the bondholders receive an extra degree of protection for which they did not pay, and the stockholders lose some dividend-paying rights for which they paid. In this way,

Table 1—*Continued*

Panel B. Firms without Documented Debt Covenant Slack Changes (FC22):

Firm	*No. of Contracts Examined*	*Remarks*
Adobe Resources	2	Both contracts used frozen GAAP.
Canadian Occidental	1	Contract contained no accounting-based covenants.
El Paso Natural Gas	2	Both contracts used frozen GAAP.
Falcon Seaboard	1	Contract used frozen GAAP.
Flying Diamond Oil Co.	1	Only accounting-based covenant was a current ratio constraint.
Husky Oil	2	One contract used frozen GAAP; other contained no accounting-based covenants.
Juniper Petroleum	1	Contract contained no accounting-based covenants.
Kaneb Services	3	Contracts contained no accounting-based covenants.
Mesa Petroleum	3	All contracts used frozen GAAP.
Natomas Co.	3	Firm reported that FAS No. 19 would have no material effect on income or retained earnings; otherwise, dividend covenant slack would have been affected.
Northwest Energy	1	Contract used frozen GAAP.
Occidental Petroleum	5	Four contracts used frozen GAAP; remaining contract limited working capital only.
Patrick Petroleum	1	Dividend slack would have been affected, but firm paid no dividends before or after FAS No. 19.
Pennzoil Co.	4	All contracts used frozen GAAP.
Rowan Cos.	3	All contracts used frozen GAAP.
Shenandoah Oil	1	Contract used frozen GAAP.
Southdown, Inc.	2	One contract used frozen GAAP; other contained leverage constraint written in terms of tangible net worth, and dividend constraint that does not depend on income.
Tenneco Inc.	5	One contract used frozen GAAP; remaining four contained no accounting-based covenants.
Tesoro Petroleum	4	One contract used frozen GAAP; one contained no accounting-based covenants; two remaining contracts contained only current ratio covenants.
Texas Oil & Gas	2	Both contracts used frozen GAAP.
Total Petroleum	1	Contract contained no accounting-based covenants.
Wainoco Oil	1	Dividend slack would have been affected, but this firm paid no dividends before or after FAS 19.

changes in the restrictiveness of dividend constraints can be costly for the stockholders of affected firms.[7]

Panel A of table 1 describes the contracts of the 13 firms that would have experienced changes in debt covenant slack. For the remaining 22 firms, the available contracts revealed no covenants that would have led to changes in slack. They are described in panel B of table 1.[8]

IV. Leverage and Stock Price Tests

On the basis of the analysis above, the 35 full-cost firms were divided into two equally-weighted portfolios: the 13 firms that would have experienced debt covenant effects if FAS No. 19 had become effective and the 22 that might not have experienced such effects. Stock price tests were performed on these two portfolios and the equally-weighted portfolio of successful-efforts firms to determine whether the 13 affected firms experienced more negative abnormal returns at the time of the Exposure Draft release and the extent to which they can explain the stock price effect observed in prior studies.

Stock Price Test of the Debt Covenant Hypothesis

Preliminary Analysis. Figure 1 shows the daily market model residuals of the full-cost and successful-efforts portfolios during the FAS No. 19 Exposure Draft period. As in Lys (1984), July 19, 1977, is the event date. As shown in figure 1, both portfolios generally exhibited negative residuals during the two-week period surrounding the Exposure Draft release. Consistent with the findings of Lev (1979) and Lys (1984), the residuals of the full-cost portfolio on the event date are more negative than those of the successful-efforts portfolio.

Figure 2 is a plot of the daily market model residuals for the successful-efforts portfolio and for the two portfolios of full-cost firms. As shown, the residuals of 13 firms with documented debt covenant effects were more negative than those of the other 22 full-cost firms. This figure suggests that the difference observed in previous studies between the returns of full-cost and successful-efforts portfolios at the time of the Exposure Draft could have been caused by a small subset of firms with debt covenant effects.

Statistical Tests. To test the significance of the differences in event date returns for these subsets of firms, I estimated the following model:

$$R_t(p)=\alpha+\beta R_{m,t}+\gamma D_t+\epsilon_t, \tag{1}$$

where:

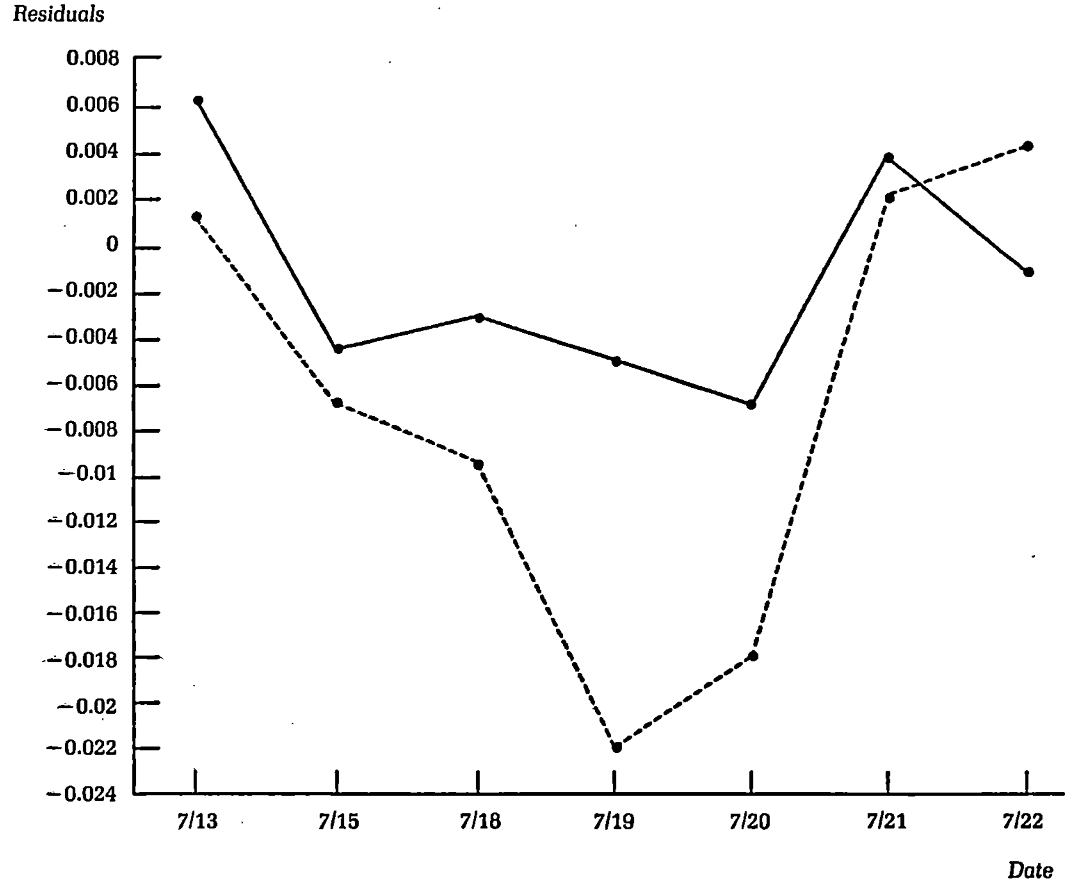
$R_t(p)$=daily returns to portfolio p, $t=1$ to 252, from January 2 to December 30, 1977;

---

[7] See Galai and Masulis (1976) and Lys (1984) for further discussion of the effect of dividend payments on the stockholders of levered firms. Kalay (1982) also describes how a binding dividend constraint can lead to overinvestment lowering the value of the entire firm.

[8] It is rarely possible to determine with certainty that a firm would not have experienced debt covenant effects from a mandated change to successful-efforts accounting. Data availability constraints make it impossible to obtain every debt contract in effect for these firms at the time of FAS No. 19. Those contracts not examined could have covenants that would have been affected. Using the available data, I can only divide firms into those with and those without documented reduction in covenant slack. The stock market and leverage tests presented here should be interpreted in light of this *caveat.*

Figure 1
Daily Market Model Residuals for the Period Surrounding July 19, 1977



Note: Residuals are based on a market model estimated over the 252 trading days from January 2 to December 30, 1977. July 14 has been omitted because the stock exchanges were closed that day because of a power failure in New York City.

——— Equally weighted portfolio of 44 successful-efforts firms
- - - - Equally weighted portfolio of 35 full-cost firms.

$R_{m,t}$ = daily returns of an equally-weighted market portfolio for the same period; and

$D_t = 1$ on July 19; 0 otherwise;

In this model, estimated on eight portfolios, the event parameter estimate, γ, isolates the abnormal return for the event date. The results are presented in table 2.

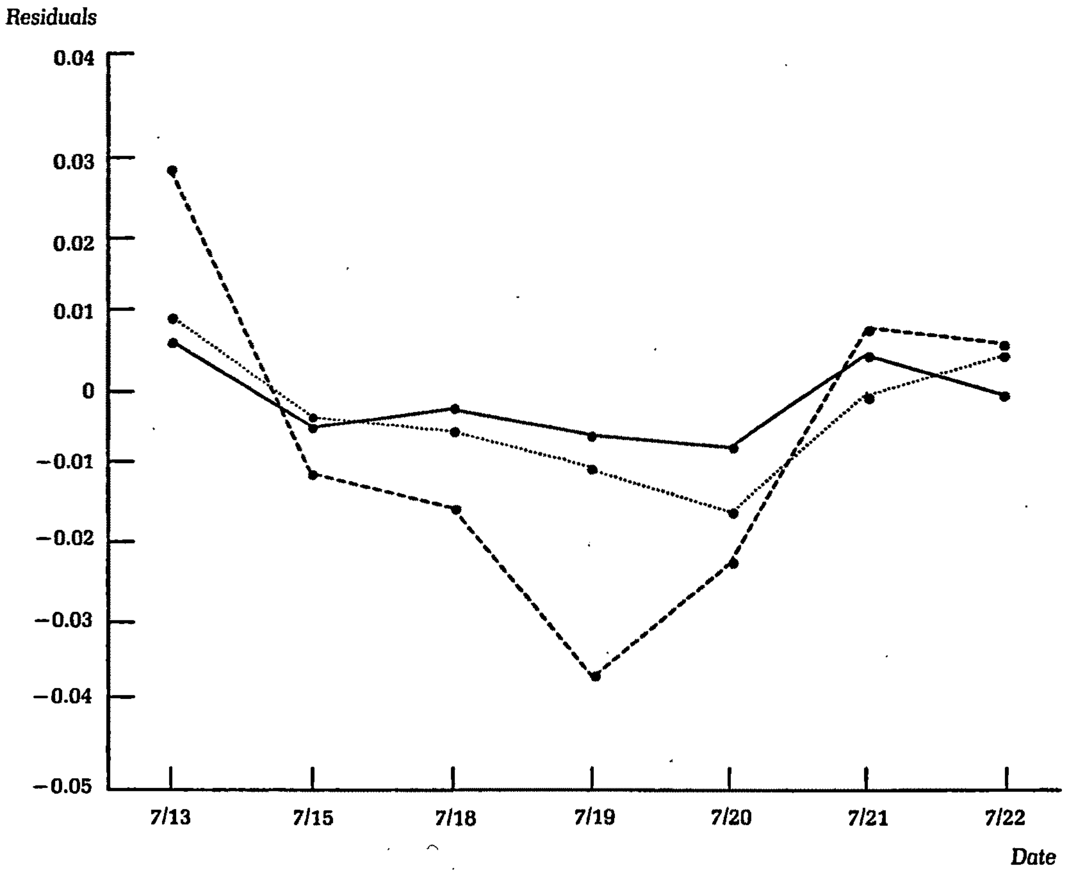
The results for the first three portfolios replicate those of Lys (1984).[9] For the port-

---

[9] Lev (1979) and Lys (1984) used an estimation period prior to the event period to estimate market model parameters in order to compute abnormal returns. I use the residual from a market model estimated over the event year, which eliminates the problem of changes in parameters between the estimation and prediction periods.

Figure 2

Daily Market Model Residuals for Portfolios of Successful-Efforts Firms and Full-Cost Firms With and Without Documented Slack Reduction for the Period Surrounding July 19, 1977



Note: Residuals are based on a market model estimated over the 252 trading days from January 2 to December 30, 1977. July 14 has been omitted because the stock exchanges were closed that day because of a power failure in New York City.

____ Equally weighted portfolio of 44 successful-efforts firms.
____ Equally weighted portfolio of 13 full-cost firms with documented slack reduction.
.... Equally weighted portfolio of 22 full-cost firms with no documented slack change.

folio of 35 full-cost firms (denoted FC), the event parameter γ was negative and significant (-0.022; $p \leq 0.006$). For the portfolio of 44 successful-efforts firms (SE), the event parameter is not expected to be significantly different from zero, since the proposed regulation affects these firms only minimally.[10] The result is consistent with this expec-

[10] None of the 44 successful-efforts firms in this sample reported that FAS No. 19 would have a materially negative effect on earnings or stockholders' equity. Still, the assumption that the event parameter for this portfolio should be zero is a strong one. Successful-efforts firms could have been affected because FAS No. 19 would have imposed a more uniform application of the successful-efforts method, or because the Exposure Draft provided evidence that all oil and gas firms were coming under closer scrutiny by regulatory bodies. All of these factors are controlled in the third portolio, which examines differences in returns between full-cost and successful-efforts firms.

Table 2
Portfolio Stock Price Regressions Based on Analysis of Contracts
(event parameter estimates and significance tests)

Model: $R_t(p)=\alpha+\beta R_{m,t}+\gamma D_t+\epsilon$

Portfolio	*Expected sign of* γ	$\hat{\gamma}$	*t*	$p < t$
FC	−	−0.022	(−3.29)	0.0006*
SE	0	−0.005	(−1.35)	0.1776**
SE-FC	+	0.017	(2.94)	0.0018*
FC13	−	−0.040	(−4.10)	0.0001*
FC22	0	−0.012	(−1.80)	0.0684**
FC22-FC13	+	0.027	(3.28)	0.0006*
SE-FC13	+	0.034	(3.88)	0.0001*
SE-FC22	0	0.007	(1.18)	0.2394**

Note:
FC = an equally-weighted portfolio of 35 full-cost firms,
SE = an equally-weighted portfolio of 44 successful-efforts firms,
SE-FC = a portfolio long in portfolio SE and short in portfolio FC,
FC13 = an equally-weighted portfolio of the 13 full-cost firms with documented slack reduction,
FC22 = an equally-weighted portfolio of the 22 full-cost firms that might have had no reduction in slack,
FC22-FC13 = a portfolio long in portfolio FC22 and short in portfolio FC13,
SE-FC13 = a portfolio long in portfolio SE and short in portfolio FC13, and
SE-FC22 = a portfolio long in portfolio SE and short in portfolio FC22.

* One-tailed test.
** Two-tailed test.

tation (−0.005; $p \leq 0.1776$). The third portfolio (SE-FC) was an equally-weighted zero investment portfolio long in portfolio SE and short in portfolio FC and provided a means of testing the significance of the difference in the impact of the event for the two groups of firms. The event parameter estimate (0.017; $p \leq 0.0018$) shows that the abnormal returns of full-cost firms were significantly lower than those of successful-efforts firms, which is consistent with Lev (1979) and Lys (1984).

The next three portfolios test whether the abnormal returns on the event date were more significant for the 13 firms with documented debt covenant effects than for the other 22 full-cost firms. The event parameter estimate for a portfolio of the 13 firms with documented covenant slack changes (FC13) was, as expected, negative and significant (−0.040; $p \leq 0.01$). The event parameter estimate for the portfolio of 22 full-cost firms with no documented debt covenant effects (FC22) was significant only at approximately the 0.07 level (two-tailed test). To test the significance of the difference in the event's effect on the two subsets of full-cost firms, the model was estimated on an equally-weighted portfolio (FC22-FC13) long in the 22 firms without documented effects and short in the 13 affected firms. The positive and significant event parameter estimate (0.027; $p \leq 0.01$) indicates that there was a significant difference between the abnormal returns of the two groups of full-cost firms on the event date.

The last two portfolios test the extent to which the firms in FC13 influence the difference between the full-cost and successful-efforts portfolios. The third test (SE-FC) was repeated twice, in each case replacing the full-cost portfolio with one of the two subset portfolios of full-cost firms. For the test using the 13 firms with covenant slack changes (SE-FC13), the event parameter doubled in size and remained significant (0.034; $p \leq 0.01$). In contrast, in the test using the portfolio of 22 firms with no documented debt covenant effects (SE-FC22), the event parameter estimate dropped from 0.017 to 0.007 and lost statistical significance ($p \leq 0.24$).

These tests offer the basis for the conclusion that firms with documented reduction in slack had a greater adverse price reaction to the FAS No. 19 Exposure Draft than any of the comparison portfolios. In my sample, these firms drive the difference between the full-cost and successful-efforts portfolios.[11]

Diagnostics

The possibility of confounding events was explored by examining *The Wall Street Journal Index* for news items about the 35 firms in question. Table 3 lists the 12 news items for the two subsamples of full-cost firms, and these are categorized as good or bad news. Only one FC13 firm had a bad news announcement. I recomputed the portfolio stock price tests after deleting that firm from the sample and found that the significance levels of the event parameters did not change.

Table 4 shows the composition of each group of firms by SIC codes. As expected, the majority of successful-efforts firms are petroleum-refining firms while the majority of full-cost firms are in the crude petroleum and natural gas category. Of the full-cost firms, the four refining firms are all in the FC22 portfolio. To test the possibility that the difference in returns between the two groups of full-cost firms may be attributable to differences in industry rather than debt contracts, I repeated the portfolio stock price tests using only firms from the crude petroleum and natural gas industry (SIC code 1311). The difference in abnormal returns between the full-cost and successful-efforts portfolios was significant at the three percent level, and the difference in abnormal returns between the two groups of full-cost firms was greater than in the original test.

Table 5 presents mean values for sales, size, leverage, beta, current ratio, and return on assets for each of the four subsets of firms in this study. In conformance with Foster's (1980) observation, the full-cost and successful-efforts groups differed significantly on every dimension except current ratio. The two groups of full-cost firms, however, showed much greater similarity. At a ten percent significance level, none of the hypotheses of equal means can be rejected. In addition, the stock price tests were repeated with a two-day event period (July 19–20) that includes the day the Exposure Draft was issued and the day it was reported in *The Wall Street Journal*. The significance levels for the event parameters did not change.

I examined event parameter estimates for the 35 individual full-cost firms to determine whether there were extreme values that could have driven the results. The fre-

---

[11] These results are consistent with some of Lev's findings (1979, 499); only 11 of the 49 full-cost firms in his sample had significantly negative returns on July 19, 1977.

Table 3
The Wall Street Journal References to Sample Firms
(July 15–20, 1977)

Content of Announcement	Date of Announcement: Firms with Documented Slack Change (FC13)	Date of Announcement: Firms without Documented Slack Change (FC22)
New chairman of subsidiary	7/15	
New debt issue	7/18	
Quarterly Income		
Good news (+)		7/18, 7/19, 7/20
Bad news (−)		7/19
Investigation of improper payments (−)	7/18	
Law suit settlement	7/15	
Oil discovery (+)	7/18	
Subsidiary closes facilities		7/15
Settled antitrust suit		7/15
Pipeline contract awarded (+)	7/19	

Note: A plus indicates an announcement expected to have a positive effect on stock prices. A minus indicates an announcement expected to have a negative effect on stock prices. A quarterly earnings announcement is considered good news if it reports earnings greater than those of the same quarter in 1976. Otherwise, it is bad news.

Table 4
Industry Composition of Portfolios

SIC Codes		Portfolios: SE	FC	FC22	FC13
1311	(crude petroleum & natural gas)	16	22	12	10
1381	(drilling oil & gas wells)	2	1	1	0
2911	(petroleum refining)	24	4	4	0
4922	(natural gas transmission)	1	8	5	3
TOTAL		43	35	22	13

Note: SE=successful-efforts firms; FC=full-cost firms; FC22=the 22 full-cost firms without documented slack reduction; and FC13=the 13 full-cost firms with documented slack reduction.

Table 5
Summary Statistics on Portfolios

	SE	FC	t-test[a] p-values	FC22	FC13	t-test[b] p-values
Sales (in $millions)	$5,338.00	$963.00	2.35 (0.02)	$1,119.00	$650.00	0.84 (0.40)
Size (in $millions)	2,824.00	731.00	2.56 (0.01)	847.00	543.00	0.97 (0.34)
Leverage	0.23	0.42	4.14 (0.01)	0.42	0.42	0.00 (0.98)
Beta	1.02	1.37	3.02 (0.01)	1.28	1.52	1.49 (0.15)
Current Ratio	1.66	1.61	0.82 (0.82)	1.31	1.76	1.56 (0.13)
Return on Assets	0.08	0.05	2.42 (0.02)	0.07	0.05	1.15 (0.26)

Note: Size is measured as the market value of common stock plus the book value of preferred stock and long-term debt, all measured at the close of 1977. Leverage is measured as the book value of long-term debt divided by size. Betas were estimated using a market model over 200 trading days from May 5, 1976 to February 23, 1977.

[a] This t-statistic tests the hypothesis that the full-cost (FC) and successful-efforts (SE) portfolios have equal means.

[b] This t-statistic tests the hypothesis that the 22 firms without documented slack reduction (FC22) and the 13 firms with documented slack reduction (FC13) have equal means.

quency distribution revealed no extreme values, which indicates that the results were not driven by one or two firms with very large stock price responses. I repeated the stock price tests, omitting the firms with the highest and lowest abnormal returns and the results were unaffected.[12]

These results are not sensitive to several alternative methods of partitioning the full-cost firms. Moving the two non-dividend-paying firms to the FC13 portfolio did not alter the results for any of the test portfolios. Similarly, moving Buttes Oil & Gas to the FC22 portfolio and making both of these changes simultaneously did not alter the results.

Tests of the Leverage Proxy

Previous tests of the debt covenant hypothesis have relied on leverage variables as proxies for debt covenant effects.[13] The analysis of debt contracts in this study provides an opportunity to test the validity of the leverage proxy by taking actual covenants and frozen GAAP features into account. The adequacy of the leverage proxy should be reflected in whether the firms most likely to be subject to debt covenant effects are more highly levered.

---

[12] For the entire full-cost portfolio, individual firm abnormal returns ranged from −0.068 to 0.021, with a mean of −0.022 and a median of 0.019. The abnormal returns for the FC13 firms ranged from −0.068 to 0.011, with a mean of −0.040 and a median of −0.046. The abnormal returns for the FC22 firms ranged from −0.053 to 0.021, with a mean of −0.012 and a median of −0.010.

[13] For examples, see Collins et al. (1981), Holthausen (1981), Leftwich (1981), and Lys (1984).

Table 6
Stock Price Regressions for Portfolios Based on Leverage

(event parameter estimates and significance tests)

Model: $R_t(p)=\alpha+\beta R_{m,t}+\gamma D_t+\epsilon$

Portfolio	*Expected sign of* γ	$\hat{\gamma}$	t	$p<t$
HI	−	−0.016	(−2.23)	0.0135*
LO	0	−0.029	(−3.53)	0.0005**
LO-HI	+	−0.012	(−1.69)	0.9455*
SE-HI	+	0.011	(1.66)	0.0495*
SE-LO	0	0.023	(3.28)	0.0012**

Note: HI = an equally-weighted portfolio of the 17 full-cost firms whose leverage was above the median leverage for all 35 full-cost firms,

LO = an equally-weighted portfolio of the 18 full-cost firms whose leverage was at or below the median leverage for all 35 full-cost firms,

LO-HI = a portfolio long in portfolio LO and short in portfolio HI,

SE-HI = a portfolio long in an equally-weighted portfolio of 44 successful efforts firms and short in portfolio HI, and

SE-LO = a portfolio long in an equally-weighted portfolio of 44 successful-efforts firms and short in portfolio LO.

* One-tailed test.

** Two-tailed test.

On average, the 13 firms with documented reduction in covenant slack were not significantly more levered than the other 22 full-cost firms. The mean leverage of these 13 firms was 0.4194, compared to the mean leverage of the remaining 22 firms of 0.4196. Additional tests were performed by dividing the 35 full-cost firms into two subgroups according to leverage. The "high" leverage group consisted of the 17 firms whose leverage was above the median, and the remaining 18 firms formed the "low" leverage group. A chi-square test of the hypothesis that grouping by leverage is independent of grouping by actual contracts cannot be rejected at conventional significance levels ($\chi^2=0.23$; $p<0.631$).

The stock price tests were repeated after replacing FC13 and FC22 with the high and low leverage portfolios, respectively. The results are shown in table 6. The first three portfolios presented in table 3 are omitted because they are unaffected by changing the partitioning of the full-cost firms. The event parameter estimate for the high leverage group (HI) is less significant than that of the low leverage portfolio (LO). The portfolio long in the low leverage firms and short in the high leverage firms (LO-HI) does not have a positive event parameter as the debt covenant hypothesis predicts. The portfolio long in successful-efforts firms and short in high leverage firms (SE-HI) does have a positive event parameter estimate, but it is significant at only the five percent level. The final portfolio, long in successful-efforts firms and short in low leverage firms (SE-LO) does have a significantly positive event parameter estimate, contrary to the prediction implied by the debt covenant hypothesis. As these results show, tests of portfolios based on analysis of actual contracts indicate that debt covenant effects can explain the stock price response to the FAS No. 19 Exposure Draft, but tests of portfolios based on leverage fail to capture that effect.[14]

V. Summary and Conclusions

Direct examination of 83 debt contracts in effect for 35 full-cost firms at the time of the FAS No. 19 Exposure Draft indicated that only 13 of the firms would have been affected by the proposed switch to the successful-efforts method. The stock price response for these 13 firms was significantly different from that of a portfolio of firms using the successful-efforts method. The stock price response of the remaining full-cost firms was not significantly different from that of the successful-efforts portfolio. The 13 firms in question did not exhibit higher average leverage ratios than the other full-cost firms, and stock price tests based on leverage did not lead to the same conclusions as those based on direct examination of contracts. These results suggest that changes in the probability of debt covenant violation are related to the stock price response to an accounting policy decision, but that a leverage variable does not provide an adequate proxy for this effect.

Appendix

The 83 debt contracts described in section IV included 29 that had provisions to protect the borrower from changes in covenant slack at the time of a mandated accounting change. This appendix includes examples of four types of provisions.

1. Some contracts "freeze" all accounting methods. This means that compliance with accounting-based covenants will be assessed on the basis of accounting procedures used by the firm in a particular set of financial statements. Mesa Petroleum's 1977 bank loan agreement states:

   > ... all of the foregoing to be determined on a consolidated basis and in conformity with generally accepted accounting principles applied on a basis consistent with the annual audit report of the Company as at the 31st day of December, 1975.

2. Some contracts freeze GAAP but allow voluntary changes within a particular set of accounting principles. Northwest Energy's 1975 loan agreement states:

   > All accounting terms not specifically defined herein shall be construed in accordance with generally accepted accounting principles in effect on the date hereof.

3. In the event of a change in accounting standards, the above two provisions would require firms to prepare accounting reports in accordance with GAAP as in effect at the date of each frozen GAAP debt contract, as well as reports in accordance with current GAAP. Some contracts have reduced the cost of producing multiple accounting reports by allowing the borrower to choose between frozen and rolling GAAP. All of Pennzoil's lending agreements refer to this provision from its 1968 Indenture:

   > Unless otherwise defined, the accounting terms used in the terms defined in this Article 1 and elsewhere in this Indenture shall be construed either in accordance with generally accepted accounting principles and practices in use at the time by the Company and its Subsidiaries or, at the option of the Company, in accordance with generally accepted accounting principles and practices in use at the date of this Indenture.

4. Another way to minimize the cost of additional accounting reports is to specify that the financial covenants will be revised in the event of an accounting rule change. Adobe Resources' 1976 Note Purchase Agreement specifies that the firm will:

   > ... keep proper books of record and account ... in accordance with generally accepted accounting principles ... except for changes in the application of accounting principles

---

[14] These results are not sensitive to the use of alternative definitions of leverage. Using each of the seven leverage definitions listed by Press and Weintrop (1990), the null hypothesis of equal means could not be rejected. The alternative leverage definitions did not affect the groupings for the other tests.

> which . . . have been consented to in writing by the holders of not less than 66⅔ percent in aggregate principal amount of the notes then outstanding *provided however*, that if such change in the application of accounting principles is mandated by statute or by the rule or regulation of any administrative agency, such written consent will not be required if the provisions of sections 9.4 to and including 9.13 shall be amended to take account of such change in application of accounting principles . . .

References

Collins, D., and W. Dent. 1979. The proposed elimination of full cost accounting in the extractive petroleum industry: An empirical assessment of the market consequences. *Journal of Accounting & Economics* 1 (March): 3–44.

——, M. Rozeff, and D. Dhaliwal. 1981. The economic determinants of the market reaction to proposed mandatory accounting changes in the oil and gas industry: A cross-sectional analysis. *Journal of Accounting & Economics* 3 (March): 37–71.

Dyckman, T., and A. Smith. 1979. Financial accounting and reporting by oil and gas producing companies: A study of information effects. *Journal of Accounting & Economics* 1 (March): 45–73.

Foster, G. 1980. Accounting policy decisions and capital market research. *Journal of Accounting & Economics* 2 (March): 29–62.

Galai, D., and R. W. Masulis. 1976. The option pricing model and the risk factor of stock. *Journal of Financial Economics* 3 (January/March): 53–79.

Holthausen, R. W. 1981. Evidence on the effect of bond covenants and management compensation contracts on the choice of accounting techniques: The case of the depreciation switch-back. *Journal of Accounting & Economics* 3 (March): 73–109.

Kalay, A. 1982. Stockholder-bondholder conflict and dividend constraints. *Journal of Financial Economics* 10 (July): 211–33.

Leftwich, R. 1981. Evidence of the impact of mandatory changes in accounting principles on corporate loan agreements. *Journal of Accounting & Economics* 3 (March): 3–36.

Lev, B. 1979. The impact of accounting regulation on the stock market: The case of oil and gas companies. *The Accounting Review* 54 (July): 485–503.

Lys, T. 1984. Mandated accounting changes and debt covenants: The case of oil and gas accounting. *Journal of Accounting & Economics* 6 (April): 39–65.

Press, E., and J. Weintrop. 1990. Accounting-based constraints in public and private debt agreements: Their association with leverage and impact on accounting choice. *Journal of Accounting & Economics* 12 (January): 65–95.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 289–303

A Perspective on Accounting-Based Debt Covenant Violations

Clifford W. Smith, Jr.
University of Rochester

SYNOPSIS AND INTRODUCTION: Some corporate decisions increase stockholder wealth while reducing the wealth of bondholders. When wealth transfers are large enough, stock prices can rise from decisions that reduce firm value. Yet, rational bondholders understand that actions taken after issuance will tend to increase stockholder wealth, and they forecast the value effects of future decision when bonds are sold. In an efficient market, the bond price at issuance reflects an unbiased forecast of the effects of such future actions. Thus, on average, bondholders will not suffer losses, although the firm (and hence its stockholders) must bear the costs of nonoptimal decisions motivated by wealth transfers from debtholders. Therefore, effective control of this bondholder-stockholder conflict can increase firm value.

Bond covenants that constrain activities such as asset sales or dividend payments are examples of voluntary contracts that can reduce the costs generated when stockholders of a levered firm follow a policy that deviates from maximization of the firm's value. The cost-reducing benefits of covenants accrue to the firm's owners through the higher price the bonds command when issued. Furthermore, if covenants lower the costs that bondholders incur in monitoring managers, these cost reductions also are passed to stockholders through higher bond prices at issuance. Therefore, in structuring an optimal debt contract, the firm's managers face a trade-off between increased proceeds from the debt issue and reduced flexibility with respect to future policy choices.

The constraints imposed through covenants are frequently specified in terms of accounting numbers. Debt covenants that employ accounting

I thank S. P. Kothari, A. Sweeney, R. Watts, and J. Weintrop for their comments. This research has been partially supported by the Simon School's Bradley Policy Research Center.

Invited by the Editor.
Received December 1992.

Editors Note: The papers by El-Gazzar and Mohrman included in this Forum were sent to Professor Clifford Smith too late to be included in his Perspective.

numbers are conventionally divided into (1) affirmative covenants, which require firms to maintain specified levels of accounting-based ratios, and (2) negative covenants, which limit certain investment and financing activities unless specified accounting-based conditions are met. For example, negative covenants restrict the payment of dividends, the disposition of assets, the issuance of additional debt, and merger activity; affirmative covenants specify minimum working capital and net worth requirements.

Although standard covenants in debt issues require that accounting numbers be consistent with generally accepted accounting principles (GAAP), they normally do not prohibit managers from switching between accepted methods. In some bank-loan agreements, firms are required only to provide unaudited, internally generated financial statements, but the contract also requires the firm to maintain substantially the same set of GAAP. If a change becomes necessary, the bank must be notified in writing prior to the change with the reasons detailed (see Zimmerman 1975).

Since different accounting techniques imply different accounting numbers, firms have incentives to relax onerous constraints through the choice of accounting techniques. Academic accountants have devoted substantial effort to obtain empirical evidence on the importance of debt agreements in determining accounting policy. (Watts and Zimmerman [1986] and Christie [1990] provide reviews.)

Initial studies adopted indirect methods to account for the effect of debt covenants on accounting decision by using the firm's debt-equity ratio as an explanatory variable in cross-sectional regressions. This ratio is a proxy for closeness to covenant constraints, as well as for the expected costs should a breach occur. Duke and Hunt (1990) and Press and Weintrop (1990) offer evidence to support the use of the debt-equity ratio as a proxy for the closeness to debt covenant constraints. Christie (1990) documents significant support for this debt hypothesis by aggregating cross-sectional studies of accounting choice, generally concluding that the larger the firm's debt-equity ratio, the more likely the firm's managers are to select accounting procedures that shift reported earnings to the current period from future periods.

Researchers have generally interpreted support for this debt hypothesis as evidence that managers act opportunistically. However, Watts and Zimmerman (1990) question whether the documented association is misinterpreted by researchers. Rather than reflecting managerial opportunism, the evidence may reflect the association among firms' investment opportunity sets, financial policies, and their efficient set of accounting methods. Even in theory, it is difficult to distinguish between opportunism and contracting efficiency as determinants of accounting policy choice. Given positive contracting costs, there will be a positive efficient amount of opportunism. Distinguishing between opportunism and efficiency is difficult in empirical work also. For example, a significant relation between accounting policy choice and leverage could indicate that managers of firms with high leverage act opportunistically in selecting accounting techniques to reduce costs imposed by constraints in debt covenants. Alternately, it could indicate that corporations for which a particular set of accounting techniques is efficient also tend to be those firms for which high leverage is efficient.

Firms examined in these cross-sectional studies are not necessarily close to their debt covenant constraints at the date examined. When firms are not close to debt covenant constraints, managerial opportunism is a less plausible explanation for the documented association between leverage and accounting choice. Yet, since it is costly for firms to switch back and forth between accounting procedures, firms that switch accounting methods to delay default are likely to continue to employ income-increasing accounting procedures, even if default is no longer likely (see Sweeney 1992). A firm's current accounting policies thus should depend on its historical choices and the time series of variables hypothesized to influence accounting policy. Therefore, cross-sectional studies do not provide the most direct or most powerful tests of the relation between accounting choice and debt contracts.

Recent studies overcome a number of limitations inherent in cross-sectional analyses by examining accounting-based defaults in debt covenants. Careful examination of the default process, its causes and cures, provides evidence on important aspects of the lending process. In this article, I review this developing literature to provide a richer understanding of the costs of leverage. These costs have important implications. For accountants, they offer potential explanations of a firm's accounting policy choice; for financial economists, they enrich our understanding of the firm's optimal capital structure.

IN section I, I focus on the role of technical default in the lending process. In section II, I examine the costs associated with contract renegotiation following default. I then analyze the implications of default costs for accounting policy choice in section III. My conclusions are presented in section IV.

I. The Technical Default Process

In the analysis of financial distress, there is an important distinction between default, bankruptcy, and liquidation. Default is the breach of a promise made under a contract. In private debt agreements, default is frequently remedied through renegotiation. Default in public debt agreements is more likely to lead to bankruptcy because of the less flexible renegotiation process and the attendant higher renegotiation costs. Bankruptcy is a legal proceeding for the court-supervised administration of a firm in financial distress. Most large public corporations file under Chapter 11, which allows reorganization of the firm's outstanding contracts. Under Chapter 7, the firm is liquidated. Liquidation is the sale of the firm's assets and distribution of the proceeds to the firm's claimants according to their priority.

It is useful to distinguish between default on a missed payment of principal or interest and the violation of other covenants, such as maintaining a specified minimum net worth or maximum leverage. Violation of a debt covenant other than one specifying promised payments is conventionally called technical default. Sweeney (1992), DeFond and Jiambalvo (1992), Beneish and Press (1993), and Chen and Wei (1993) focus on technical defaults; they analyze firms that announce initial breaches in an accounting-based debt covenant. Current accounting standards mandate disclosure of debt covenant defaults. For example, SFAS No. 59 specifies that default is a basis for a going-concern warning, and SFAS No. 78 requires disclosure of a breach when long-term debt is reclassified as a current liability. SEC rules also require that any violation that exists at the date of a balance-sheet filing must be disclosed in the notes to the finanical statements. Thus, by examining public documents, the authors identify cases of default.

Their analyses document important facts about technical defaults. First, virtually all the technical defaults occur in private rather than public debt issues. This is consistent with the hypothesis of Smith and Warner (1979) and Leftwich (1981) that, because contract renegotiation is less costly in private debt issues, covenant constraints will be set with less slack, and hence default will be more frequent than in public issues. Sweeney (1992) suggests that technical defaults in public issues frequently result from cross-default provisions in the public debt contracts and thus are rarely incremental defaults in the sample.

Second, the covenants most frequently violated are (in decreasing order): net worth (or tangible net worth), working capital (or current ratio), leverage, interest coverage, and cash flow (or cash flow coverage). The evidence thus indicates that firms more frequently violate affirmative covenants; the breach of negative covenants is rare. The violation of an affirmative covenant gives lenders additional rights, such as to accelerate maturity of the debt, but lenders do not have the right to call a loan simply because a negative covenant is binding. For default to occur in connection with a negative covenant, managers must pay dividends or issue new debt when debt covenants restrict such actions.

Third, the evidence suggests that firms frequently violate multiple covenants: 49 percent of the default firms violate one covenant, 32 percent violate two covenants, 18 percent violate three covenants, and 1 percent violate four covenants. Two factors are potentially important in explaining multiple covenant violations: accounting measures on which covenant constraints are based are not independent; a firm with deteriorating performance will have that fact related across several dimensions. Also, because SFAS No. 78 requires long-term debt be reclassified as a current liability unless a waiver in excess of one year is obtained, there is a mechanical association between other unwaived defaults and the breach of a working-capital covenant.

Fourth, the available evidence suggests that technical defaults are generally triggered by the deterioration of firms' operating performance rather than by mandated changes in GAAP. Beneish and Press (1993) report no technical defaults triggered by mandated accounting changes, and Frost and Bernard (1989) report only two. Several factors are potentially important in explaining this low frequency. First, as El-Gazzar (1993) emphasizes, firms can alleviate the effects of mandated accounting changes by renegotiating potentially binding agreements and by retiring debt. Second, changing GAAP is frequently a lengthy process. The longer the time from an initial discussion memorandum to final implementation, the more time firms have to adjust and the less likely the changes will result in technical default. Third, when changes in GAAP do cause a material change in a covenant-imposed constraint, private lenders have strong incentives to renegotiate debt contracts. For the lender, maintaining a reputation for eschewing opportunistic behavior is important in attracting future borrowers and, thus, is a valuable asset. Finally, the power of the methods used by Beneish and Press (1993) to identify technical defaults triggered by change in GAAP is limited. Mandated accounting changes adopted by their sample firms within their sample period typically increased earnings per share. More data are required to better determine relative frequencies.

The Waiver Decision

The lender's option to accelerate debt maturity can be waived or employed to force contract renegotiation. The articles by Beneish and Press (1993) and Chen and Wei (1993) examine the creditor's decision to waive events of technical default.

Beneish and Press (1993) identify 91 firms that violate accounting-based covenants between 1983 and 1987. They divide the sample into (1) firms that receive waivers without alteration in contract terms (12 firms), (2) firms that receive waivers following contract renegotiation (53 firms), (3) firms that fail to obtain a waiver from existing lenders; but obtain financing from new lenders (8 firms), and (4) firms that receive no waiver but do not renegotiate the contract within one year of default (18 firms).

Beneish and Press (1993) document interesting differences among these subsamples. Compared to firms that receive waivers without contract renegotiation, the firms that receive waivers following renegotiation have significantly larger increases in interest rates, have significantly more covenants added to the agreement, and receive significantly lower abnormal returns around the announcement of the technical default. Compared to firms that successfully renegotiate, the firms that change lenders after failing to receive waivers have significantly greater increases in interest rates, have significantly more new covenants added to the agreement, and receive significantly lower announcement-date abnormal returns. Finally, compared to the firms that receive waivers following renegotiation, the firms with no waivers and no renegotiation also receive significantly lower abnormal returns at the announcement of default.

The overall evidence suggests that lenders deal with different firms in technical default differently. This observation suggests that firms in the various classifications are likely to face materially different circumstances and, thus, there is a potentially important selection bias in the data. Care must be taken to control for these differences before drawing inferences. For example, the abnormal return evidence indicates that the market distinguishes among these firms at the date that the firm's financial statement reporting the default is released. But technical default is only one of the pieces of information reflected in these measured abnormal returns. The abnormal returns reflect both the incremental costs of technical default as well as the incremental value implications of the other information disclosed in the firm's financial statement. Because technical defaults reflect deteriorating operating performance, this other information is likely to comprise a significant component of the measured abnormal return.

Chen and Wei (1993) identify 128 firms that disclose the occurrence of technical default in their annual reports between 1985 and 1988. They identify whether a waiver is granted by examining management's discussion and analysis and footnotes. In their sample, 57 firms (45 percent) receive a waiver after the violation; 71 (55 percent) do not. Of the 57 waivers, 24 are for a limited period and 33 are permanent.

The grouping of firms by whether a waiver is granted is an issue that deserves careful attention. Standard loan documentation gives the lender the option to accelerate the maturity of the loan if technical default occurs, but the contract does not require acceleration. (Beneish and Press [1993] report 18 firms that do not receive a waiver yet do not renegotiate the lending agreement.) The firm receives an implicit waiver so long as the lender does not exercise the option to accelerate the debt maturity. This argument suggests that, rather than focusing on whether an explicit waiver is granted (as in Chen and Wei's [1993] Models I and II in table 5), the more appropriate concern is whether the lender exercises the option to accelerate maturity and force renegotiation (as in Model IV of table 5). This view implies that the lender's reactions to default fall along a continuum: at one extreme, the lender grants a permanent waiver without renegotiation; next, a temporary waiver without renegotiation; then, no waiver, but no renegotiation; then, a waiver following renegotiation; finally, renegotiation fails, no waiver is granted, and the firm obtains financing from another lender. Note that this is not a debate about where to draw a line along a continuum, but a suggestion that the initial Chen and Wei analysis groups dissimilar events.

Using the intuition of the option-pricing model, Chen and Wei (1993) suggest that waivers are more likely to be granted in some cases than others. For example, they argue that a firm with a higher probability of bankruptcy is less likely to receive a waiver because the opportunity cost to the lender of losing the borrower is lower. Thus, Chen and Wei argue that a waiver is more likely: the lower a firm leverage, the shorter the maturity of the issue, the lower the principal amount of the issue, and if the issue is secured. Employing LOGIT regression techniques, they examine the association between the waiver decision and leverage, maturity, principal amount, and security provisions. With the exception of debt maturity, their estimated coefficients are significant with the predicted sign.

Chen and Wei (1993) address two related issues: the lender's waiver decision and the debt hypothesis for accounting policy choice. Yet, they are not always clear about the differences between the two issues. There is a subtle problem with their LOGIT regression analysis of the waiver decision. They state that the intuition behind the effect of leverage is that the lower the leverage ratio, the better the debt is protected. Hence, the creditors have less reason to call the debt. This intuition would be sufficient if firms were homogeneous and covenants were uniform across debt issues. But the constraints negotiated between borrowers and lenders vary across debt agreements.

Smith and Watts (1992) argue that an important determinant of the firm's financial-policy choice is its investment opportunity set. With risky fixed claims included in the firm's capital structure, benefits from a profitable investment project can go to the bondholders to a degree that the cash flows received by the firm's stockholders do not offer them a normal rate of return. Stockholders thus can have incentives to reject a project with a positive net present value. Myers (1977) calls this the underinvestment problem. Firms whose value depends more on their subsequent investment decisions face higher underinvestment costs and thus will employ less leverage. Conversely, firms whose value is composed primarily of tangible assets will employ more leverage.

This argument implies that a pharmaceutical company with a large budget for research and development (R&D) would employ less leverage than a gas pipeline company. Therefore, a debt agreement for the pipeline would allow higher leverage than would one for the drug company. If both contracts were optimally set and both leverage covenants were breached, it is not obvious that a lender would be more likely to waive the violation for the drug company than for the pipeline. Even though it has less leverage, the pharmaceutical firm faces higher costs from the disincentive effects of debt. Therefore, the waiver decision should depend not on the level of leverage, but on the deviation from the firm's optimal leverage. For example, suppose the typical pharmaceutical company had 10 percent leverage while the typical gas pipeline had 50 percent. A lender reasonably might be more concerned about a drug company with 35 percent leverage than a pipeline with 55 percent.

The LOGIT regressions of Chen and Wei (1993) thus suffer from an omitted-variables bias because there is no control for the firm's investment opportunity set. This problem affects several of the variables: (1) The firm's measured leverage will be correlated with its optimal leverage and hence with its investment opportunity set. (2) Myers (1977) argues that the firm's debt-maturity choice will be driven by its investment opportunity set; Barclay and Smith (1992) document that a firm with more growth options chooses shorter maturity debt. (3) Security interests are generally granted on tangible rather than intangible assets. Therefore, the firm's investment opportunity set will be correlated in Chen and Wei with the variables measuring leverage, debt maturity, and the existence of security interests. I am also concerned that the measure of issue size in Chen and Wei proxies for firm size since larger firms make larger debt issues. Therefore, it is difficult to interpret their estimated coefficients as tests of their hypotheses to explain the lender's waiver decision.

Despite the problems with this LOGIT regression in explaining the waiver decision, I believe that their results can be interpreted as useful evidence on the debt hypothesis. By documenting that waivers are more likely the lower the firm's leverage, they provide evidence supporting the assumption that the level of debt is related to the expected costs of default.

Temporary Versus Permanent Waivers

Chen and Wei (1993) analyze the lender's decision with respect to the duration of waivers of technical defaults. It is important to understand the range of actions available to a lender facing a technical default; by focusing attention on temporary versus permanent waivers, Chen and Wei document important texture in the lending process. They eliminate permanent waivers and employ LOGIT regression techniques to examine the association between the choice of temporary waivers and leverage, maturity, principal amount, and security provisions.

I am not surprised at the authors' limited success in explaining the choice of temporary versus permanent waivers. Notwithstanding the above-mentioned specification problems, I expect that lenders base this decision on more information than the authors examined. Temporary waivers should be more frequently granted when the technical violation is triggered by a nonrecurring event. If the borrower explains that the problem producing the default is temporary, then granting a temporary waiver (or granting an implicit waiver by not accelerating maturity and not forcing renegotiation) can be a prudent lending decision. If the borrower's explanation is accurate, then the lender's interests are unimpaired and forcing more dramatic action is inappropriate. Note that in this case the omitted variables are unlikely to be correlated with the included variables. Thus, although the regression potentially lacks power, the reported coefficients should be unbiased.

II. Cost of Technical Default

Interest Rate Changes

Beneish and Press (1993) document a change in borrowing terms for 48 of their 91 firms that report technical default. Rates increase for 31 of the 48 firms, decrease for 3 firms, and remain unchanged for 14 firms. The mean change in rates is 80 basis points; the median is 55. To provide an estimate of the cash-flow effect of this change, they apply the rate change to the outstanding loan amount in the year after the violation. The mean estimated incremental interest cost for one year is $293,800; the median is $85,300. As a fraction of the firms' equity, the mean is 0.0095; the median is 0.004. Note that these numbers probably include a small downward bias. Over the sample period (1983–1987), interest rates generally were falling. Since most loan agreements include prepayment options, the four firms in their sample with fixed-rate loans would normally renegotiate a lower rate. The estimates of refunding costs by Beneish and Press thus potentially understate the rate changes expected if market interest rates were constant.

Beneish and Press (1993) then estimate a cross-sectional regression to explain the variation in the ratio of their estimated incremental interest cost to the market value of the firm's equity. The independent variables they employ are the firm's leverage, return on assets, the rate of return on the firm's shares, and a dummy variable to indicate if the debt is secured. Of their estimated coefficients, only leverage is significant.

Like Chen and Wei (1993), Beneish and Press (1993) are interested in two different but related questions. They analyze changes in interest rates on loans to explain the renegotiation process, and provide additional evidence on the debt hypothesis for accounting policy choice. However, for explaining the renegotiation process, their interest-cost regression is misspecified. There is a basic inconsistency between the left and right sides of their regression equation. To explain the change in interest cost to the firm following default requires information, not on the level of leverage in the firm's capital structure, but on the change in leverage.

There are at least three ways to address this problem of dimensional inconstancy: (1) add the leverage at the time the firm originally negotiated the loan terms; (2) add measures of the firms' investment opportunity set as independent variables (as in Chen and Wei's [1993] waiver regression); here one must assume the appropriate change in rates is related to the deviation in leverage from that of the firm's optimal capital structure; and (3) employ a model using a portfolio of accounting measures to derive an instrument for the probability of bankruptcy (see, e.g., Ohlson 1980), then employ the estimated change in the probability of bankruptcy as an independent variable.

The Press and Weintrop (1990) analysis suggests a fourth potential approach: add the percentage deviation from the constraint specified in the covenant as an independent variable. But note that this last approach is likely to be difficult to implement when contracts and the particular covenants violated vary across sample firms. For example, if a contract does not specify a maximum leverage, one could not employ the difference between leverage and the constraint as an independent variable. If one tried to employ the percentage deviation from the particular constraints violated, another problem arises. It is not clear which covenant should be analyzed if a firm violates multiple constraints. If one simply chooses the largest deviation, the underlying distribution of the error terms for the regressions is unlikely to be homoskedastic. Without using generalized least squares, the tests would be inefficient, although the estimated coefficients would be unbiased.

Covenant Modifications

Beneish and Press (1993) examine the modification of covenants following technical defaults. They examine covenants constraining leverage, tangible net worth, working capital, the current ratio, and retained earnings. Of the 163 covenants they observe, 56 are relaxed and 107 are tightened or unchanged. They calculate the percentage change in the level of the constraint. In their sample, the mean change relaxes the constraint for each type of covenant with the exception of those restricting retained earnings.

I believe that Beneish and Press (1993) would derive a richer explanation of these changes if they segregated the renegotiated covenants by whether technical default is triggered by a breach of that covenant. I expect that most of the covenants in which constraints are relaxed are those that have been breached. Conversely, the covenants that are tightened should be other covenants. The observations by Beneish and Press that covenants restricting retained earnings and dividends are frequently made more restrictive are consistent with this conjecture since they report that covenants specifying minimum earnings account for less than 4 percent of their violations and dividend covenants are never the source of default in their sample.

Beneish and Press (1993) document the addition of 184 covenants for the 61 firms that renegotiated their debt agreements following technical violation. For these firms, the mean percentage change in the number of covenants is 26.6; the median is 33.3. Less than 10 percent of the new covenants are accounting-based. This statistic potentially reflects the widespread use of accounting-based covenants in these agreements (obviously they cannot be added if they are already present). The additional covenants generally specify security interests in assets; limit capital expenditures, acquisitions, stock repurchases, and additional debt; or require prior approval for new financing. Thus, the added covenants are typically negative covenants.

Beneish and Press (1993) employ regression analysis to examine the covenant additions. They regress the percentage change in the number of covenants in the renegotiated agreement on leverage, the return on assets in the default year, the stock return in the default year, and a dummy variable to indicate whether the debt is secured. They find only the return on assets is significant.

This regression suffers from the same specification bias as that discussed with respect to their analysis of interest-rate changes. However, this analysis also raises another problem. The changes in interest rates, changes in covenant constraints, and the addition of new covenants are substitutes. In the renegotiation process, if a borrower agrees to more covenant-imposed constraints on future activities, one would expect that any increase in interest rates would be smaller. However, great care must be exercised in the empirical analysis to reflect this interdependence among the variables appropriately. Across the sample, the circumstances of the firms are not homogeneous; thus, the observed variation in these dimensions typically will be positively correlated. Some firms will have unimportant covenant breaches and will receive waivers without contract renegotiation; other firms will suffer major breaches and will receive both higher interest rates and more binding constraints on future activities in renegotiation. However, this is like observing that, although gold and diamonds are substitutes, wealthy individuals tend to have a lot of both, and low-income individuals tend to have few of either.

Default Costs and Lender Characteristics

In addition to the borrower characteristics focused on in the analyses by Beneish and Press (1993) and Chen and Wei (1993), lender characteristics are also potentially important in explaining variation in the costs of technical default. Lending policies differ. Suppose that at origination one lender sets a leverage constraint 1 percent below its customer's current leverage, while another sets it 10 percent below. If both borrowers breach the covenant by 2 percent, the lenders are likely to react quite differently. The first should be more willing to waive the violation with only a 3 percent leverage change; the second should be more likely to accelerate maturity with a 12 percent change. There is an important difference in negotiating over a covenant breach with one bank that makes a $100 million loan versus 20 banks that loan $5 million each. The larger the group of lenders providing the credit, the more likely are conflicts among lenders, the greater the free-rider problems among the lenders, and the more likely is a lender to adopt a hold-up strategy. Therefore, I expect higher renegotiation costs with more lenders in the loan syndicate. (Note that Chen and Wei report that 39 percent of their sample involves private debt with multiple creditors.)

Regulation of lenders can introduce variation in default costs. For example, a bank that is receiving additional regulatory scrutiny because it has violated capital guidelines is likely to exhibit less flexibility in dealing with a borrower in technical default. Therefore, the financial condition of the lender can be important. Regulation also can influence details of the lender's response to default; for example, many firms in technical default have unused portions of revolving credit agreements. Upon default, lenders often reduce allowed borrowings under these agreements. In the risk-based capital guidelines applied to banks, unused portions of irrevocable commitments enter risk-weighted asset calculations at 100 percent. Unused commitments with a maturity of more than one year have a conversion factor of 50 percent, and unused commitments with a maturity of less than one year have a conversion factor of zero. This regulation provides incentives for banks to decrease allowed borrowings under commitments that are irrevocable, as well as those that have maturities of greater than one year. However, for unused commitments with maturities of less than one year, this banking regulation introduces no additional incentive to limit further borrowings.

Omission of variables to control for lender characteristics reduces power, but does not introduce bias in the estimated coefficients so long as the omitted lender characteristics are uncorrelated with the included independent variables. This condition is likely to be met for the regulation variable. Unfortunately, at our current level of understanding of the lending process, we do not know the extent of the correlation for the other lender characteristics. For example, a particularly risky loan customer might be more likely to go to a single lender to internalize any incentive problems among lenders. Alternatively, such a loan customer might go to a consortium of lenders because of lender risk-management stategies. Thus, it is difficult to assess the sign or magnitude of the potential bias this omission introduces in the reported regressions.

Policy Changes Following Default

Beneish and Press (1993) examine changes in the default firms' investment, financing, and payout policies within one year of violation. They find in their sample of 91 firms that 54 divest assets, 12 reduce dividends, 12 reduce R&D expenditures, nine privately place common stock with major shareholders, eight issue warrants or refinance and exchange debt, six cut their labor forces, and five terminate pension plans. It is critical to remember that these transactions are not all the result of technical default. If the firm were in exactly the same business circumstances except that the constraints in the debt covenants were not breached, many of these same policy choices would be taken. For example, R&D expenditures are a class of discretionary investment activities that are particularly susceptible to reduction when a firm suffers a loss in value. This facet of the analysis by Beneish and Press therefore provides useful evidence on the broader question of changes in corporate policies associated with financial distress.

Beneish and Press (1993) focus particular attention on eight firms that issue warrants to lenders or issue a new package of debt and stock in exchange for the old debt. I also find these cases interesting. When a lender is dealing with a financially distressed firm, incentive problems can become extremely difficult. With traditional loan products, the lender takes most of the downside while the borrower gets most of the upside. The promised interest rate necessary to offer the lender a normal expected rate of return can be quite high in such circumstances. By restructuring through the provision of warrants to the lender, for example, the lender shares in the upside, and thus the conflict is better controlled. Moreover, since warrants make the lender's compensation contingent on future firm value, their use reduces the borrower's liquidity problems.

III. Default Costs and Accounting Policy Choice

Default and Abnormal Accruals

DeFond and Jiambalvo (1992) and DeAngelo et al. (1992) examine the association between debt covenants and accounting accruals. DeFond and Jiambalvo estimate abnormal accruals for 94 firms that report covenant violations in their annual reports. They find that abnormal accruals are significantly positive in the year prior to violation. In the year of violation, they document evidence of positive abnormal accruals after controlling for management changes and auditor going-concern qualifications.

However, the power of any test that employs abnormal accounting accruals is limited by the methods available to estimate the abnormal accruals. These methods potentially misclassify accounting changes that involve changes in cash. For example, DeAngelo et al. (1992) classify inventory reductions as income-decreasing. However, if the inventories liquidated are LIFO inventories, then cash and reported earnings can increase. Hence, such LIFO liquidations are potentially misclassified as income-decreasing by typical abnormal accrual measures. Correct classification of abnormal accruals is also difficult because of the negative serial correlation of changes in accruals (see Dechow 1992), as well as the correlation between abnormal accruals and changes in the levels of firms' economic activities. This problem is likely to be especially severe in a sample of financially distresssed firms. To reduce this problem, DeFond and Jiambalvo (1992) employ cross-sectional and time-series controls.

Default and Accounting Technique Choice

To avoid the problems with documenting discretionary changes in accruals, researchers have examined changes in accounting techniques. Beneish and Press (1993) examine defaulting firms' accounting policies with respect to depreciation, inventory, pension costs, and investment tax credits. They report that defaulting firms employ income-increasing accounting policies with significantly greater frequency than that reported by Zmijewski and Hagerman (1981) or Press and Weintrop (1990), both one and five years prior to the technical default. As Sweeney (1992) notes, using the results of other studies as a benchmark requires care. Systematic misclassification of firm's accounting choices is more likely in a sample of firms approaching technical default than for a random sample of firms, as in Zmijewski and Hagerman. For example, if a firm approaching technical default is in a declining-cost industry, then FIFO is income-decreasing. Also, through its effect on work-in-process and finished-goods inventories, straight-line depreciation can decrease working capital. Thus, switching to accelerated depreciation or FIFO can tighten working-capital constraints.

To examine whether managers change accounting procedures in response to tightening constraints, Sweeney (1992) examines the time series of managers' accounting choices before and after firms violate debt covenants. She documents that managers of firms approaching such constraints are more likely to make income-increasing discretionary accounting changes than a sample of control firms matched by industry, size, and time period. Managers of default firms make a larger number of income-increasing changes in the violation year than in surrounding years. Moreover, managers continue to make income-increasing changes in years following the first year of default. Relative to managers of control firms, managers of default firms also accelerate the adoption of mandatory accounting changes that increase reported earnings.

Sweeney (1992) documents that managers of 130 default firms made 205 accounting changes in the period from five years prior to two years following the year of technical default and 76 percent of these changes were income-increasing. Although this evidence is consistent with the hypothesis that managers make these changes to offset increases in the tightness of debt covenant constraints, the evidence could also reflect changes in optimal accounting policies due to the firms' changed circumstances. For example, if prior to default, the firm's operating cash flows decline, the reduced liquidity could motivate the firm to liquidate LIFO layers, terminate overfunded pension plans, and change pension-accounting assumptions. These accounting changes also might be motivated by concomitant changes in tax circumstances.

To help sort out the importance of these alternate motives for the observed accounting changes, Sweeney (1992) provides a detailed examination of 22 default firms that disclose the boundaries of the covenants that they violate. She finds that accounting changes delay technical default for five of the firms. In four additional cases, managers could have delayed default by switching inventory-valuation methods, yet they did not because switching from LIFO to FIFO would have imposed significant tax costs. Managers thus appear to balance the benefits of avoiding default against the tax costs of changing inventory-valuation methods.

Sweeney (1992) notes that even substantial default costs will not motivate accounting policy changes under some circumstances. For example, when managers' accounting flexibility is too small to avoid the violation, the firm is not expected to change. Evidence from her cross-sectional analysis of defaulting firms' accounting policy choices is consistent with this hypothesis.

Negative Covenants and Accounting Choice

The evidence from the research on technical default helps in understanding the results of Healy and Palepu (1990) and DeAngelo et al. (1992). Healy and Palepu examine firms for which the dividend constraint becomes more binding. They conclude that firms cut dividend payments but do not make accounting changes to circumvent the dividend restrictions. DeAngelo et al. examine accounting choices for firms with persistent losses and dividend reductions. They find that approximately 40 percent of the reductions are forced by binding dividend constraints. Their evidence suggests that managers' accounting choices primarily reflect acknowledgement of their firms' financial difficulties via negative abnormal accruals and discriminatory write-offs, rather than attempts to inflate reported income. Thus, while the dividend covenant appears to be an important constraint in determining the firm's payout policy, affirmative covenants appear more important than negative ones, such as the dividend covenant, in explaining accounting policy choice.

IV. Conclusions

Examination of the evidence on technical defaults yields a much richer understanding of the lending process. The evidence suggests that technical default generally is caused by deteriorating corporate performance, that it almost exclusively involves affirmative covenants, that a substantial fraction of the defaults involve the breach of multiple constraints, that default is more likely in private than public debt issues, and that lenders choose among a wide array of measures to respond to a technical default. We are beginning to understand the choice of actions by the borrower and lender facing technical default. For the lender, these actions range from granting a permanent waiver without further modification of the contract to accelerating the maturity of the debt and forcing the borrower to obtain financing elsewhere. The possibility of technical default affects the borrower's accounting, financing, payout, and investment policies in potentially important ways. Moreover, borrowers face policy choices both in anticipation of the possibility of default as well as in response to default, including negotiations with lenders.

It is critical that researchers more consistently model the activities they examine in order to make more progress. In the model, mathematics is not necessary, but it is important for reseachers to carefully analyze the incentives of various market participants making the decisions. Such analysis offers the only reliable guide to the factors that should be examined as explanatory variables as well as the specification of the estimation equation. More care in this regard would have eliminated a number of the specification problems discussed here.

We also must develop a clearer perspective of the relative frequency of technical default and default on promised debt payments across public as well as private debt issues, the frequency of renegotiation versus bankruptcy, and the frequency of reorganization versus liquidation. It is important to analyze conditional probabilities; for example, given an initial technical default in a private debt agreement, what is the probability of default on a promised payment within a particular period? The probability of contract renegotiation? The probability of bankruptcy? The probability of liquidation?

I believe that further progress requires the integration of technical default into a broader view of the lending process. For example, the focus on initial technical defaults in present studies limits our understanding in potentially important ways. Suppose the manager of a portfolio of loans regularly sets constraints just below the firms' current values. For most firms in the loan portfolio, business is normal or improving, covenants are not breached, promised payments are made on a timely basis, and required reports are delivered. But for a subset of the firms in the loan portfolio, results deteriorate, technical default occurs, and the borrower must contact the lender under standard contracts. After obtaining the facts, including the borrower's forecast of future performance, the lender chooses an action. Assume that the lender waives the default but resets the constraint, again just below the firm's current level. If the borrower's operating performance improves, the lender just cashes checks and files reports. But if results deteriorate further, the borrower must contact the lender again. This lending strategy produces a dynamic interaction between borrower and lender that provides continuing flexible monitoring of loan customers, while focusing on those borrowers that are most likely to generate losses.

Although we have much still to learn about the role of default, bankruptcy, and liquidation in the determination of the firm's optimal policy choices, research to date provides a productive beginning.

References

Barclay, M., and C. Smith. 1992. The maturity structure of corporate debt. Unpublished manuscript, University of Rochester, NY.

Beneish, M., and E. Press. 1993. Cost of technical violation of accounting-based debt covenants. *The Accounting Review* 68 (April): 233–57.

Chen, K., and K. Wei. 1993. Creditors' decisions to waive violations of accounting-based debt covenants. *The Accounting Review* 68 (April): 218–32.

Christie, A. 1990. Aggregation of test statistics: An evaluation of the evidence on contracting and size hypotheses. *Journal of Accounting Economics* 12: 15–36.

DeAngelo, H., L. DeAngelo, and D. Skinner. 1992. Accounting choice in troubled companies. Unpublished manuscript, University of Michigan, Ann Arbor.

Dechow, P. 1992. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. Unpublished manuscript, University of Rochester, NY.

DeFond, M., and J. Jiambalvo. 1992. Debt covenant violation and manipulation of accruals. *Journal of Accounting & Economics* (Forthcoming).

Duke, J., and H. Hunt III. 1990. An empirical examination of debt covenant restrictions and accounting-related debt proxies. *Journal of Accounting & Economics* 12: 45–63.

El-Gazzar, S. M. 1993. Stock market effects of the closeness to debt covenant restrictions resulting for capitalization of leases. *The Accounting Review* 68 (April): 258–72.

Frost, C., and V. Bernard. 1989. The role of debt covenants in assessing the economic consequences of limiting capitalization of exploration costs. *The Accounting Review* 64 (October): 788–808.

Healy, P., and K. Palepu. 1990. The effectiveness of accounting-based dividend covenants. *Journal of Accounting & Economics* 12: 97–123.

Leftwich, R. 1981. Evidence of the impact of mandatory changes in accounting principles in corporate loan agreements. *Journal of Accounting & Economics* 3 (1): 3–36.

Myers, S. 1977. Determinants of corporate borrowing. *Journal of Financial Econnomics* 5 (2): 147–75.

Ohlson, J. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* (Spring): 109–31.

Press, E., and J. Weintrop. 1990. Accounting-based constraints in public and private debt agreements: Their association with leverage and impact on accounting choice. *Journal of Accounting & Economics* 12: 65–95.

Smith, C., and J. Warner. 1979. On financial contracting: An analysis of bond covenants. *Journal of Financial Economics* 7 (2): 117–61.

——, and R. Watts. 1992. The investment opportunity set and corporate financing dividend and compensation policies. *Journal of Financial Economics* 32 (3): 263–92.

Sweeney, A. 1992. Debt-covenant violations and managers' accounting strategies. Unpublished manuscript, Harvard University, Cambridge, MA.

Watts, R., and J. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice-Hall.

——, and ——. 1990. *Positive accounting theory: A ten-year perspective. The Accounting Review* 65 (January): 131–56.

Zimmerman, C. 1975. An approach to writing loan agreement covenants. *Journal of Commercial Bank Lending*: 2–17.

Zmijewski, M., and R. Hagerman. 1981. An income strategy approach to the positive theory of accounting standards setting choice. *Journal of Accounting & Economics* 3 (2): 129–49.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 304–322

Evidence on the Nature of Audit Planning Problem Representations: An Examination of Auditor Free Recalls

Mary York Christ
The Pennsylvania State University

SYNOPSIS AND INTRODUCTION: Audit planning is a crucial task in the audit process. While making planning decisions, auditors have a large amount of information available to them and their existing knowledge structures influence how this information is used to form a problem representation. This problem representation has a significant effect on all subsequent decision making. The purpose of this study is to investigate the nature and evolution of auditors' planning problem representations by an examination of free recalls.

In the cognitive psychology literature, it has been reported that, as decision makers accumulate more experience, their knowledge structures change to embody more total knowledge, more knowledge of relationships, and more abstract knowledge. In the context of an audit planning task, I predicted that problem representations formed by more experienced auditors would reflect these characteristics. To test the hypotheses, an experimental study was conducted with 211 practicing auditors. Subjects were from 16 offices of the then Big Eight public accounting firms, with experience levels ranging from one month to 30 years. Using experimental materials that consisted of approximately 20 pages of background information on a hypothetical audit client, subjects completed a planning task, followed by a background questionnaire, and a surprise recall task. The free

This article is based on my dissertation. I would like to thank my committee members for their help and suggestions while completing this research: L. Tomassini, U. Anderson, B. May, and D. Schkade. This article has benefited greatly by the thoughtful comments of the referees and, especially, the associate editor. Helpful comments on earlier drafts were received from M. Abdolmohammadi, R. Christ, B. Cushing, M. Dirsmith, J. Hansa, J. McKeown, R. Tubbs, I. Vessey, and J. Webster. Financial assistance from the Deloitte, Haskins & Sells Foundation, the Institute of Internal Auditors Research Foundation—A. O. Hinkle Memorial Fund and The University of Texas at Austin are gratefully acknowledged.

Submitted May 1991
Accepted November 1992.

recalls provide indirect evidence on the nature of the problem representations formed during the planning task.

The results suggest that the problem representations of managers and partners may be different from those of seniors or juniors. The manager/partner recalls included a smaller percentage of simple case facts and a higher number and percentage of abstractions from the basic facts. These results have potential implications for audit practice. To the extent that the managers and partners form "better" problem representations and are actively involved in the process, more effective and efficient audit planning might be possible. However, more research needs to be conducted to determine the specific effect of the exhibited differences on planning decisions.

Key Words: *Audit planning, Experience, Free recall, Problem representation.*

Data Availability: *The data upon which this research is based may be obtained from the author on request.*

PLANNING is a crucial step in the audit process. In the initial planning phase, auditors develop expectations about the likelihood of errors in the financial statements and design an audit strategy that is appropriate for the circumstances (Felix and Kinney 1982). Their expectations are formed as a result of considering a variety of different variables, including information from previous audit working papers, minutes of meetings of the board of directors, interim financial statements, preliminary analytical review procedures, discussions with the client, trade and industry journals, and relevant economic and environmental issues. These variables are interrelated, and the organization of the auditor's knowledge about these factors is crucial for conducting an effective audit (Felix and Kinney 1982).

Given the large amount of information available to the auditor during audit planning, how that information is interpreted is an important issue. Cognitive psychology research suggests that existing *knowledge structures* stored in memory influence how available information is combined to form a *problem representation* of the decision situation. This problem representation significantly affects the efficiency and effectiveness of all subsequent information processing (Glaser 1984); hence, it should have a significant effect on audit planning decisions. Moreover, the literature on experience and expertise suggests that knowledge structures change as experience increases (Chi et al. 1982); hence, audit planning problem representations may be affected by experience. Ashton et al. (1988, 108) note the importance of these constructs[1] to auditing as follows:

> The role of schemata is particularly relevant in auditing, because the schemata developed by auditors through experience and prior knowledge of client situations may affect the manner in which the auditor perceives the evaluation of assertions and the need to accumulate and interpret evidence about these assertions.

For example, consider an auditor planning the procedures to be performed regarding the valuation assertion for accounts receivable. The auditor might note that accounts receivable, the allowance for doubtful accounts, and bad debt expense are all comparable to last year. However, the auditor may fail to relate that observation to the fact that the general health of the industry has deteriorated this year and many of the client's customers are facing financial problems. The problem representation formed might include a conclusion that collectibility is not a major issue, with minimal procedures being planned in this area. If in fact collectibility is a problem and accounts receivable are overstated, then the effectiveness of the audit has been impaired because of this aspect of the auditor's problem representation. Similarly, scenarios could be conceived where inadequate problem representations lead to inefficiencies in allocating audit effort.

---

[1] Ashton et al. (1988) use the term *schemata* which can be defined as knowledge structures stored in memory that represent generic concepts (Rumelhart 1984).

The purpose of this research is to investigate the nature of auditors' planning problem representations and to examine how this construct changes with experience. An experimental study was conducted in which practicing auditors with different levels of experience completed an audit planning task followed by a surprise free recall task. In this latter task, subjects were asked to recall everything they could remember about the case. The recalls provide indirect evidence on the auditors' problem representations.

This study contributes to the literature by examining an aspect of auditor knowledge for which very little reseach has been conducted (Bonner and Pennington 1991). Previous research has examined auditor knowledge of errors (see, e.g., Ashton 1991; Libby 1985; Libby and Frederick 1990; Tubbs 1992) and the organization of auditor knowledge of internal controls (see, e.g., Frederick 1991; Weber 1980). These prior studies provide important insights into auditor domain knowledge, but the current study focuses on how that knowledge is brought to bear on the problem representation in audit planning.

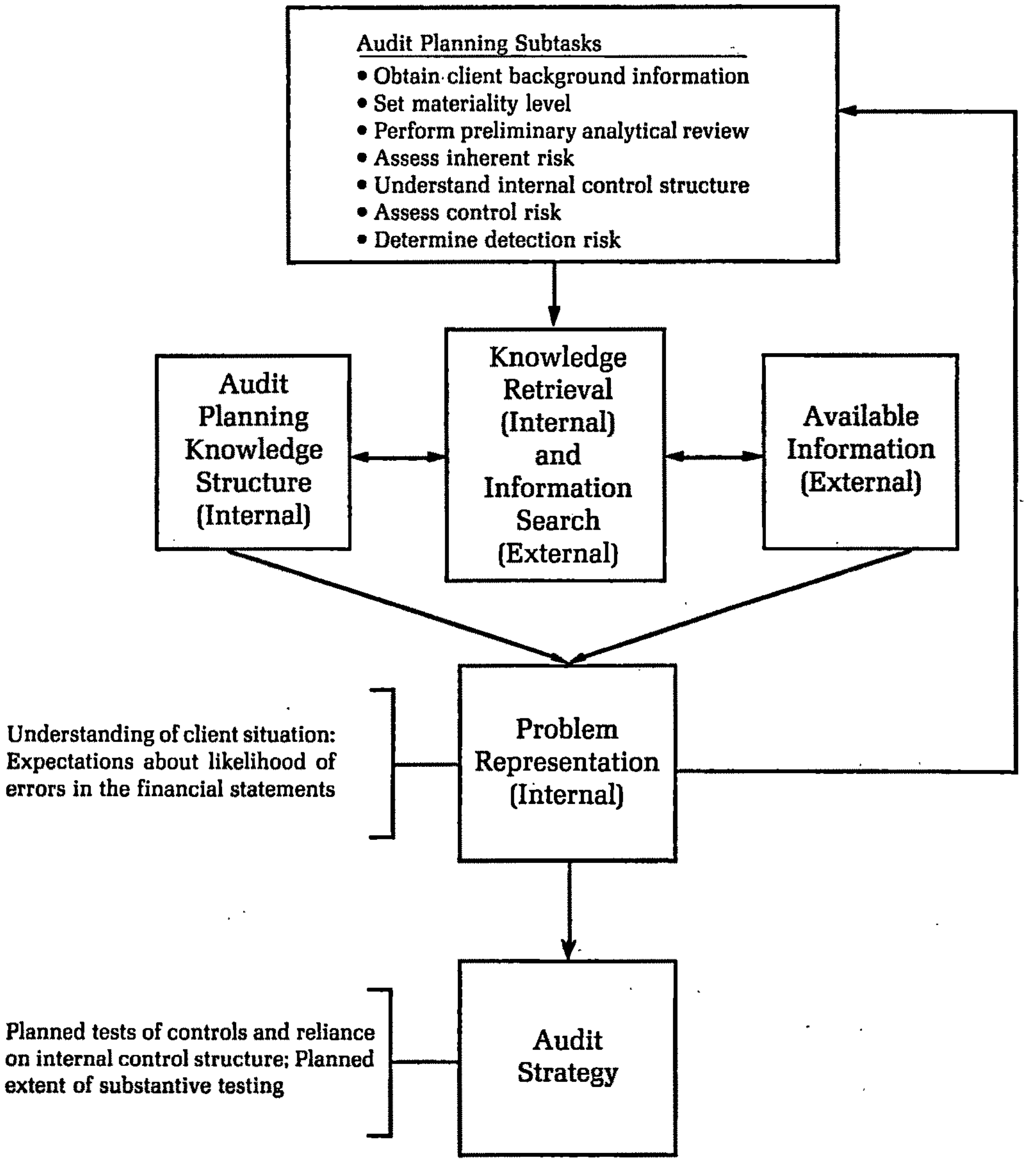
The remainder of this article is organized as follows. The theoretical background for the study is discussed in section I, and the experimental method used is explained in section II. The results are presented in section III, and the results and implications of the research are discussed in section IV.

I. Theoretical Background

The Audit Planning Problem Representation

The ultimate goal of the planning task is to develop an audit strategy appropriate for the circumstances. This strategy is based on the auditor's expectations about the likelihood of errors in the financial statements (Felix and Kinney 1982). In cognitive terms, these expectations are embodied in the auditor's problem representation. A problem representation is a decision maker's understanding and interpretation of a problem situation (Chi et al. 1981) and is developed by mapping the information available for completing a task into an existing knowledge structure related to that type of task. The knowledge structure represents the current state of the decision maker's domain task knowledge. The knowledge structure's organization and content have a significant effect on what information is searched for and how the information is organized and interpreted in the problem representation (Glaser 1984). Figure 1 provides a simplified cognitive model of the audit planning process. As the various planning subtasks are completed, the problem representation evolves and the auditor develops an understanding of the likelihood of errors in the financial statements. Throughout this

Figure 1
The Audit Planning Task


Audit Planning Subtasks
• Obtain client background information
• Set materiality level
• Perform preliminary analytical review
• Assess inherent risk
• Understand internal control structure
• Assess control risk
• Determine detection risk
Audit Planning Knowledge Structure (Internal)
Knowledge Retrieval (Internal) and Information Search (External)
Available Information (External)
Understanding of client situation: Expectations about likelihood of errors in the financial statements
Problem Representation (Internal)
Planned tests of controls and reliance on internal control structure; Planned extent of substantive testing
Audit Strategy

process, audit planning knowledge guides the search for relevant information and the mapping of that information into the problem representation. The problem representation then becomes an integration of the available information and the auditor's knowledge about the planning task.

Previous research has demonstrated the significance of the problem representation for subsequent problem solving (Glaser 1984). For example, Carroll et al. (1980) found that providing subjects with a graphic problem representation significantly improved their performance on a design problem. They suggest that the representation aided the subjects in dealing with an otherwise difficult and ill-structured problem. In the context of medical diagnosis, Kassirer and Kopelman (1987, 1989) illustrated the importance of the problem representation for a correct diagnosis. In auditing, Johnson et al. (1991) show that the correct representation is critical for an audit reporting decision, and Bedard and Biggs (1991) found that identifying a particular pattern of cues as part of the problem representation was a necessary condition for determining the cause of an unexpected fluctuation in an analytical review task.

Experience and Knowledge Structures

Domain-related knowledge develops as a result of experience (Bonner and Pennington 1991; Chi et al. 1982). Consequently, as auditors acquire more experience, their planning knowledge structures should become more developed and their planning problem representations may also change. By the time auditors reach the senior level, they have had training in most planning subtasks and are at the experience level generally considered necessary for planning (Bonner and Pennington 1991). Although experience alone does not ensure that an auditor will become "expert" in a particular task (Bonner and Lewis 1990; Davis and Solomon 1989), it does provide an opportunity for the development of expertise. Moreover, the expertise literature generally recognizes more developed knowledge structures as a necessary component of expertise.

Previous research has generated a robust set of findings on how the knowledge of experts differs from that of novices. This literature, therefore, provides a useful theoretical basis for hypothesizing how experience might affect the development of audit planning knowledge, which in turn affects the formation of the problem representation.[2] In general, research has shown that more developed knowledge structures embody (1) more knowledge, (2) more knowledge of *relationships*, and (3) more *abstract* knowledge than do less developed structures (Chi et al. 1982). Each of these three characteristics is discussed in turn.

Total Knowledge. As individuals accumulate more experience in a domain, they acquire an increasing body of knowledge. This increased knowledge may come from additional formal training as well as from specific experiences in the domain (Bonner and Lewis 1990; Chi et al. 1982). In previous audit studies, Weber (1980), Frederick and Libby (1986), Libby and Frederick (1990), Ashton (1991), and Tubbs (1992) all found that more experienced auditors had more knowledge.[3] This increased knowledge should influence the problem representation formed.

---

[2] Although the expertise literature provides part of the theoretical background, this is not a study of expertise. Rather, the purpose of this research is to examine differences in problem representation between experience levels with an ultimate goal (in future research) of assessing how any differences identified relate to task performance and expertise (Davis and Solomon 1989). It is also worth noting that some of the expertise literature has classified experts and novices on grounds other than task performance. For example, Chi et al. (1982) classified two physics professors with "considerable experience teaching introductory physics" as experts and freshman physics majors as novices.

[3] Ashton, however, found that experienced auditors had increased knowledge about only the most frequently occurring financial statement errors.

H1: As experience increases, the audit planning problem representations that auditors construct will reflect a mapping of the available information into a knowledge structure that contains more total knowledge.

Knowledge of Relationships. Individuals with more experience develop more knowledge of how different variables are related to each other (Chi et al. 1982). This type of knowledge allows them to recognize patterns of information from a set of cues. One implication of this relational knowledge is that individuals store incoming information in "chunks" or groupings that are meaningful for the particular domain (Glaser and Chi 1988; Lesgold et al. 1988; McKeithen et al. 1981).

In auditing, knowledge of how various patterns of cues relate to the likelihood of errors in the financial statements is important for effective planning. For example, among the audit planning subtasks listed in figure 1, knowledge of relationships is important for analytical review (cf. Bedard and Biggs 1991) and for the evaluation of the internal control structure. In previous research, Frederick and Libby (1986) found that experienced auditors had knowledge of the relations between control weaknesses and account errors and knowledge of relations among accounts, whereas inexperienced auditors had only the latter knowledge. Frederick (1991) found that experienced auditors' recalls of internal controls were both greater in quantity and more tightly clustered when presented to them in a schematic versus a taxonomic organization.[4] In a workpaper review task, Moeckel (1990) found that less experienced auditors did not integrate information (link inconsistent cues) as often as more experienced auditors. Increasing knowledge of relationships should influence the problem representation formed.

H2: As experience increases, the audit planning problem representations that auditors construct will reflect a mapping of the available information into a knowledge structure that contains more knowledge of relationships.

Abstract Knowledge. The development of abstract knowledge involves learning the deeper principles associated with a particular domain. In other words, as individuals become more experienced, they learn to *abstract* from the basic facts to more overriding concepts, which allows them to make inferences that go beyond the information given. One manifestation of more abstract knowledge is the finding that novices tend to focus on the surface features of a problem, whereas experts organize their problem representations around the functional features or "deep structure" of the problem.[5] Expertise researchers have found this surface/functional distinction in a variety of domains, including physics (Chi et al. 1982) and computer programming (McKeithen et al. 1981).

Abstract knowledge may be particularly important for identifying major issues during audit planning. For example, classifying a client as having a going-concern problem or a high risk of management fraud may involve abstracting from the total set of cues. These types of conclusions could have a pervasive effect on all planning decisions.

---

[4] In general, taxonomic organizations are based only on principles of class inclusion, whereas schematic organizations are based on spatial or temporal *relationships* between members (Mandler 1984).

[5] Surface features are words or items specifically referred to in the problem, or key words that have a particular meaning in the domain. In contrast, the deep structure relates to basic principles or rules that govern the solution of the problem.

Biggs et al. (1988) conducted a verbal protocol study involving an analytical review and audit program design task. Their results indicate that senior auditors were oriented toward obtaining only the information necessary to complete the task. In contrast, experienced managers spent more time getting a thorough understanding of the client firm and the nature of its business and problems. In other words, the less experienced auditors looked only at the surface details of the case and the more experienced ones searched for overriding issues relevant to the client. Bouwman (1982) found a similar result in a verbal protocol study of expert and novice financial analysts.

In other research, Sandlin (1987) found that inexperienced auditors were not as sensitive to changes in the background field as more experienced ones when explaining an unexpected fluctuation during analytical review. This finding is consistent with experienced auditors' going beyond the surface features of a problem. Finally, Moeckel (1990) found that, when reviewing workpapers, experienced auditors reconstructed information (i.e., abstracted and made inferences) more than less experienced auditors. Increasing abstract knowledge should also influence the problem representation formed.

H3: As experience increases, the audit planning problem representations that auditors construct will reflect a mapping of the available information into a knowledge structure that contains more abstract knowledge.

II. Method

Subjects

To test the hypotheses, an experimental study was conducted. Auditors at three different levels of experience completed a planning task followed by a surprise recall task. Subjects were 211 auditors from a total of 16 offices of the Big Eight[6] public accounting firms in two major southwestern cities and one major midwestern city. Their experience levels ranged from one month to 30 years, with an average of 4.5 years. The auditors were classified into three experience levels: (1) junior staff, (2) senior staff, and (3) managers/partners. Discussions with practicing auditors suggest that the primary responsibility for planning generally falls upon the senior auditor. Furthermore, Bonner and Pennington (1991) note that for most of the subtasks involved in audit planning, approximately 3.5 years of experience are required (with the exception of evaluating the internal control structure, which requires approximately 1.5 years of experience). The experience groupings in this study, therefore, allow for a comparison of those auditors who typically perform the planning task to groups of both less experienced and more experienced auditors. The distribution of subjects across experience levels is as follows:

---

[6] The data for this study was collected prior to the mergers within the Big Eight firms.

	Years of Experience	
	Average	*Median*
Juniors (n=82)	1.0	1.0
Seniors (n=55)	3.5	3.3
Managers/partners (n=74)	9.3	7.7
Total (n=211)	4.5	3.0

Materials and Procedure

The materials for the experiment consisted of a set of general instructions, the case booklet, and a series of tasks. The case booklet was approximately 20 pages long and provided a detailed description of the industry and company, including a narrative description, an organizational chart, document and systems flowcharts for the revenue cycle, an internal control questionnaire for the revenue cycle, comparative income statements and balance sheets, and comparative accounts receivable distributions and agings.

In task 1 of the experiment, the auditors were asked to make a series of planning judgments related to the revenue cycle. The purpose of task 1 was to encourage the auditors to process the information in a "planning mode" and subjects were free to review the case booklet at any time while completing this task. Information processing is goal-directed (Rumelhart 1984), so it was important that the auditors think the only purpose of the research was to study audit planning judgments. Task 2 was a background questionnaire given as a distracter task prior to the third task, a surprise free recall. In task 3, subjects were given ten minutes to recall everything they could about the case material.[7] They were not allowed to look back at the case booklet or the previous tasks.

The experiment was conducted in the offices of participating firms, with from nine to 22 auditors completing the study in each office. The subjects proceeded through the materials in one sitting, in the presence of the researcher, and returned the materials before leaving. On average, it took them approximately one hour to complete the entire experiment.

Operationalization of Hypotheses

The auditors' surprise free recalls provide indirect evidence on their problem representations as well as the knowledge structures accessed during the planning process.

---

[7] The ten-minute time limit was based on results of a pilot study and appeared to be an adequate, but not excessive, amount of time.

First, the recalls provide evidence about the problem representations because the auditors were asked to recall *case-specific* information rather than only generic auditing knowledge.[8] The recalls should reflect their interpretation of the case materials. Furthermore, since a problem representation is a mapping of the available information into an existing knowledge structure, the recalls also provide some indirect evidence on the nature of the knowledge structures accessed.[9]

The use of a recall methodology is supported by some of the basic findings related to the nature and functions of knowledge structures (Brewer and Nakamura 1984). First, information is better recalled if a subject has an existing knowlege structure in which to map the information rather than no knowledge structure (Bransford and Johnson 1972; Chiesi et al. 1979). Second, subjects with a more developed knowledge structure in a particular domain will show higher recall for materials related to that knowledge (Chiesi et al. 1979). Finally, information related to a knowledge structure will be recalled better than unrelated information (Brewer and Treyens 1981; Mandler and Johnson 1977; Rumelhart 1977).

In this study, the recall protocols were first segregated into a series of complete thoughts or "idea units" (Newell and Simon 1972; Bransford and Johnson 1972). The idea units were then coded in a manner similar to techniques used in verbal protocol studies, by focusing on the character of the information recalled. Although this approach has not been used extensively, it is similar to some of the analysis techniques used by Chi et al. (1982) on subject summaries and elaborations and by Pennington (1987) on computer program summaries and recalls.

The idea units were classified into one of the following categories: (1) basic facts (FAC), direct statements or simple paraphrasing of information specifically stated in the case; (2) relationships (REL), statements that relate information in the case; this could include summarizing two or more items of information of the same type or relating two or more items of information of a different type; and (3) abstractions (ABS), statements that include information not presented in the case; this could involve adding an evaluative component, such as an adjective or adverb, to case information or going beyond the case information by suggesting explanations of the case material, asking questions about missing information, or making statements about auditing the company. Examples of each category are given in the appendix.

The three hyspostheses predict, respectively, that as experience increases, auditors' problem representations will reflect a mapping of the available information into knowl-

---

[8] This approach is different from that in some auditing recall studies. For example, in Libby and Frederick (1990), auditors recalled generic auditing knowledge about financial statement errors, which provided evidence on their knowledge structures. In another study, Frederick (1991) asked subjects to recall a list of internal controls he had presented to them. However, when the subjects received the list, they were told that they would be asked to recall the items later. Although they may have mapped the control information into an existing knowledge structure, they did not do so in an audit decision-making context. Therefore, Frederick's results also focus on knowledge structures as opposed to problem representations. The approach in the current study is similar to that of Choo and Trotman (1991). Their subjects reviewed a description of a client with a potential going-concern problem and then later completed a surprise recall task. However, Choo and Trotman focused their analysis and discussion on inferences about the auditors' knowledge structure instead of their problem representations.

[9] Note that this study does not attempt to infer the relationship between information recalled and how the information was actually used in planning decisions. The ultimate decision involves weighting and combining the information, as well as a consideration of risk preferences—issues that are beyond the scope of this study. However, there is evidence that information recall (vs. information recognition) is consistent with the decision made (see, e.g., Dellarosa and Bourne 1984; Norman et al. 1989).

Table 1
Summary of Hypotheses and Operationalization

Hypothesis		*Prediction*	*Operationalization in Free Recall*
H1	⇑Experience⇒	⇑Total Knowledge Reflected in Problem Representation	⇑Total Recall
H2	⇑Experience⇒	⇑Knowledge of Relationships Reflected in Problem Representation	⇑REL Recall
H3	⇑Experience⇒	⇑Abstract Knowledge Reflected in Problem Representation	⇓FAC Recall ⇑ABS Recall

Note: FAC = basic facts; REL = relationships; and ABS = abstractions.

edge structures containing more total knowledge, more relational knowledge, and more abstract knowledge. The construct related to total knowledge (hyp. H1) was operationalized by examining the total number of idea units recalled by the subjects. The construct dealing with relational knowledge (hyp. H2) was tested by examining the number of REL idea units recalled. As discussed, knowledge of relationships involves recognizing patterns of information and relating different variables to each other, including chunking of the information. Therefore, hypothesis H2 predicts an increase in the REL category with experience.

Finally, the construct for abstract knowledge (hyp. H3) was tested by looking at the number of FAC and ABS idea units recalled. As a direct restatement of case facts, FAC idea units represent the surface information in the case. According to hypothesis H3, more experienced auditors will recall fewer FAC idea units. Abstract statements go beyond the basic case in some way, so statements such as "has strong internal controls" or "involved in a volatile industry," which include an evaluation of the case information, were coded as ABS.[10] Some ABS statements go even further beyond the basic facts of the case, often asking questions or commenting on how to audit the company. Examples include "obsolete technology a problem for realizability" or "credit problems?" (see the app.). All the ABS statements involve abstracting from the basic facts, perhaps to overriding principles the auditor considers important for planning the audit. Hypothesis H3, then, predicts an increase in ABS statements with experience.[11] Table 1 summarizes the operational constructs used to test each of the hypotheses.

The recalls were coded independently by the reserher and a paid rater who was unaware of the experimental hypotheses (Newell and Simon 1972). Both coders determined the number of idea units per recall and the classification of the idea units. The intraclass correlation (Guilford and Fruchter 1978, 268–70) for the two counts of total idea units per subject was 0.895. Differences were discussed and resolved to make the total idea units per subject the same for both raters. Agreement on the FAC, REL, and ABS classifications, as measured by the kappa coefficient (Bishop et al. 1975, 395–96), was 81.7 percent. Remaining differences in classification were resolved by the researcher.

---

[10] The case materials included information on controls but never stated that they were "strong." Similarly, the materials described the industry, but did not specifically identify it as "volatile."

[11] It should be noted that no attempt was made to assess the *quality* of the ABS or REL statements.

III. Results

Firm Structure Effects

Recent research on audit firm methodologies has revealed that firms differ in their audit approaches (Cushing and Loebbecke 1986; Kinney 1986). Since subjects from all of the Big Eight firms participated in this study, the structure of the firms' audit methodology could have a confounding effect on the results. Using the structure categories reported by Kinney (1986), subjects were classified according to whether they were from structured firms ($n=62$), semistructured firms ($n=95$), or unstructured firms ($n=54$). To test for structure effects, this classification was included as an independent variable in all of the initial data analyses. Firm structure was not statistically significant, either as a main effect or as an interaction, and was subsequently dropped from the analyses.

Tests of Hypotheses

If hypothesis H1 were true, more experienced auditors should recall more idea units. A one-way analysis of variance on the total number of idea units recalled by each experience level revealed a significant experience effect [$F(2,208)=12.39$; $p=0.0001$]. Because the group differences were in the expected direction, Tukey tests of paired comparisons were conducted to determine which differences were statistically significant. The results generally support hypothesis H1; however the difference in total recall between seniors and managers/partners is only marginally significant. The mean number of idea units recalled and the standard deviations for each group are as follows:

Juniors (18.5; 5.6 s.d.) vs. Seniors (21.1; 6.6 s.d.)	$p \leq 0.05$
Juniors (18.5; 5.6 s.d.) vs. Managers/partners (23.5; 6.6 s.d.)	$p \leq 0.001$
Seniors (21.1; 6.6 s.d.) vs. Managers/partners (23.5; 6.6 s.d.)	$p \leq 0.10$

Table 2 summarizes the auditors' recalls by category and experience level. Hypotheses H2 and H3 require that subsets of these recall categories be examined. However, the various categories are likely to be correlated, so a multivariate analysis of variance (MANOVA) was performed first to see if experience had any overall effect on the results. The MANOVA, which used the number of idea units recalled in each category as the dependent variables, revealed a significant experience effect [$F(6,412)=15.60$; $p=0.0001$]. Bonferroni multiple comparisons were subsequently used to test hypotheses H2 and H3. The results are reported in table 3.[12]

Hypothesis H2 predicts an increase in REL recall as experience increases. As shown in table 3, the only significant difference in REL recall was between juniors and the manager/partner group. However, given the results of the tests for hypothesis H1, it is possible that any significant differences found for the individual categories may be confounded with total recall differences. This possibility was evaluated in two ways. First, the total frequency of idea units recalled in each category for each experience

---

[12] Since the hypotheses are directional, the results in table 3 are based on one-tailed Bonferroni comparisons. A question may be raised about the appropriateness of following a two-tailed MANOVA with one-tailed comparisons. If the Bonferroni tests were made on a two-tailed basis, the difference between juniors and seniors on the number of ABS statements recalled would be significant at the 10 percent level instead of the 5 percent level reported. All other results would remain the same.

Table 2
Idea Units Recalled

	Categories			
Experience Level	*FAC*	*REL*	*ABS*	*Total*
Juniors				
Average number recalled	13.4	2.2	2.9	18.5
Standard deviation	5.5	2.1	2.2	5.6
Average percentage of total recall	72.5	12.0	15.5	100.0
Seniors				
Average number recalled	14.3	2.5	4.3	21.1
Standard deviation	7.5	2.1	3.4	6.6
Average percentage of total recall	67.9	11.6	20.5	100.0
Managers/partners				
Average number recalled	12.7	3.2	7.6	23.5
Standard deviation	6.7	2.3	3.9	6.6
Average percentage of total recall	54.2	13.5	32.3	100.0
Total Subjects				
Average number recalled	13.4	2.6	4.9	20.9
Standard deviation	6.5	2.2	3.8	6.6
Average percentage of total recall	64.1	12.5	23.4	100.0

Note: FAC = basic facts; REL = relationships; and ABS = abstractions.

level was computed. A chi-square test for dependence between experience level and category ($\chi^2 = 150.5$, $p = 0.0000$) indicates that the differences cannot be explained by amount of recall. Second, a MANOVA was performed on the arcsin transformation (Winer 1971, 400) of the proportion of idea units recalled in each category. Experience is still a highly significant factor [$F(6,412) = 6.75$, $p = 0.0001$]. However, subsequent Bonferroni comparisons, shown in table 3, reveal no significant differences related to the percentage of total recall accounted for by REL statements. In total, these results provide only very limited support for hypothesis H2, with respect to differences between juniors and managers/partners.[13]

Hypothesis H3 predicts a decrease in FAC and an increase in ABS recall with experience. As seen in table 3, when the raw number of statements recalled is compared, there are no FAC differences. However, for the ABS category and amount of recall, there are significant differences between all three of the groups. When the proportion of idea units in each category is analyzed, both the FAC and the ABS differences are significant for the managers/partners versus each of the other two groups. The FAC findings can be explained as follows. With increasing experience, the managers/partners did not recall a significantly smaller *number* of FAC statements. However, to the extent that they recalled more items, the additional recall was in other categories, caus-

---

[13] The REL category may be further subdivided by the type of REL statements made. Specifically, some REL statements such as "Bad debts are up" involve grouping and summarizing *similar* pieces of information (SIM), while others such as "A/R up in spite of lower sales" relate *different* types of information to each other (DIF). Bonferroni paired comparisons were also made for these subcategories. These tests revealed no significant differences for the SIM statements, but the managers and partners recalled both a greater number and higher percentage ($p \leq 0.05$) of DIF statements than either of the other two groups. The total number of DIF statements was quite small, however. Across all subjects, the average number was 0.4, which on average constituted 2.2 percent of total recall.

Table 3
Bonferroni Paired Comparisons

	Significant Differences in Recall	
Experience Levels	*Idea Units Recalled*	*Percentage of Total*
Juniors vs. Seniors	ABS*	
Juniors vs. Managers/partners	REL*, ABS**	FAC**, ABS**
Seniors vs. Managers/partners	ABS**	FAC**, ABS**

Note: FAC=basic facts; REL=relationships; and ABS=abstractions.

* Group difference significant with $p \leq 0.05$.
** Group difference significant with $p \leq 0.01$.

ing the *proportion* of FAC recall to decrease. Overall, these results support hypothesis H3 for differences between the manager/partner group and the other two levels and provide limited support for differences between juniors and seniors.

Supplemental Analysis

The analyses presented thus far have focused on the character of the recalls. As knowledge structures develop, it is also possible that auditors will focus on different *content* areas. For example, analytical review procedures have been shown to be diagnostic and useful in directing attention during planning. Therefore, as knowledge of financial relationships increases, the problem representations formed might focus more on financial information than on other types of information. A second analysis was performed in which the content of the recalls was considered. There is no strong theoretical basis for predicting how the content of planning knowledge structures will change with experience. Therefore, the content coding provides supplemental and exploratory evidence about the auditors' problem representations.

For this coding, the idea units were classified according to the following scheme: (1) statements about any of the basic client and industry information in the case were classified as general client (GC); (2) statements about the information provided in the comparative financial statements or accounts receivable agings and distributions were classified as financial (FS); (3) statements about the general system design and the internal control procedures were classified as internal control (IC); and (4) statements about information not in the case were classified as other (OT).[14] Examples of each of these categories are presented in the appendix.

The recalls were coded for content by using the same procedure as for the coding by character. The intraclass correlation for the two raters' counts of total idea units per subject was 0.921.[15] The kappa coefficient measuring agreement on the classifications was 93.9 percent and the remaining differences were resolved by the researcher.

---

[14] The OT category for content is very similar to the ABS category for character.

[15] The number of idea units for the character and content codings was not always the same. For the content codings, an idea unit was one piece of information. For the character coding, an idea unit might involve linking two pieces of information together. For example, the statement "Experiencing decreasing sales with no decrease in COGS" would be one REL idea unit in the character coding but two FS idea units in the content coding.

Table 4
Recall Coded by Content

Experience Level	*Categories*				
	GC	*FS*	*IC*	*OT*	*Total*
Juniors					
Average number recalled	13.6	3.3	2.2	0.6	19.7
Standard deviation	5.4	2.7	2.1	1.1	6.2
Average percentage of total recall	68.8	16.9	11.3	3.0	100.0
Seniors					
Average number recalled	14.8	5.0	2.0	1.1	22.9
Standard deviation	6.5	3.3	1.8	1.5	7.0
Average percentage of total recall	64,5	21.7	9.0	4.8	100.0
Managers/partners					
Average number recalled	14.0	6.5	2.8	2.6	25.9
Standard deviation	5.8	3.2	2.4	3.0	7.3
Average percentage of total recall	54.0	25.2	10.7	10.2	100.0
Total Subjects					
Average number recalled	14.0	4.9	2.4	1.4	22.7
Standard deviation	5.8	3.3	2.1	2.2	7.3
Average percentage of total recall	61.8	21.5	10.4	6.3	100.0

Note: GC=general client; FS=financial information; IC =internal control; and OT=other.

Table 4 summarizes the results of the content coding. The results of a MANOVA performed on the number of idea units recalled in all the content categories indicate a significant experience effect [$F(8,410)=11.50$; $p=0.0001$].[16] Subsequent Bonferroni tests of paired comparisons are presented in table 5.

The paired comparisons of the number of idea units recalled indicate significant differences between all groups for the FS category and between the manager/partner group and each of the other two groups for the OT category. A chi-square test for dependence between experience level and category ($\chi^2=143.9$, $p=0.0000$) indicates that the differences cannot be explained by amount of recall. A MANOVA on the arcsin transformation of the proportion of idea units recalled in each category also indicates that experience is a highly significant factor [$F(8,410)=7.43$, $p=0.0001$]. The Bonferroni comparisons of proportions reported in table 5 show that (1) the proportion of GC recall decreased with each experience level, (2) juniors recalled a smaller proportion of FS statements than the other two groups, and (3) managers/partners recalled a larger proportion of OT statements than the other two groups. Because the OT content category is fairly similar to the ABS character category, the main insight gained from this analysis is related to the GC and FS results.

---

[16] There were two versions of the case materials, and half the subjects completed each version. Case version was not a significant factor in the total recall or character coding analyses and was dropped from those analyses. In the content coding analysis, case version is a significant factor ($p=0.0111$). However, when it is included in the analysis, the experience factor still remains highly significant ($p=0.0001$), and the conclusions of the study do not change. There was no significant interaction between case version and experience.

Table 5
Bonferroni Paired Comparisons of Content Recall

	Significant Differences in Recall	
Experience Levels	*Idea Units Recalled*	*Percentage of Total*
Juniors vs. Seniors	FS*	GC*, FS*
Juniors vs. Managers/partners	FS**, OT**	GC**, FS**, OT**
Seniors vs. Managers/partners	FS*, OT**	GC**, OT**

Note: GC = general client; FS = financial information; IC = internal control; and OT = other.

* Group difference significant with $p \leq 0.05$.
** Group difference significant with $p \leq 0.01$.

IV. Discussion

The purpose of this study was to increase our knowledge of the planning problem representations of auditors with different levels of experience. The results provide mixed support for the hypotheses. First, with respect to total recall, managers/partners and seniors recalled significantly more case information than juniors, and managers/partners recalled marginally significantly more information than seniors. These results suggest that the problem representations of different experience levels do reflect different total amounts of information and that the knowledge structures that influenced those problem representations may include different amounts of planning knowledge. Although the recalls of managers/partners included only marginally more idea units than those of seniors, the FAC and ABS recall results suggest other differences between their problem representations.

With respect to the knowledge of relationships, the results provide only very minimal support for any differences. The only significant result found was in the number of REL statements made by juniors versus managers/partners.[17] Finally, with respect to abstract knowledge, the results support differences between managers/partners and the other two levels but provide only limited support for any differences between seniors and juniors. With respect to the latter two groups, the only difference identified was in the number of ABS statements made. Managers/partners, however, recalled a smaller proportion of direct case facts (FAC) and both a greater number and proportion of abstractions (ABS) than did juniors or seniors.

The supplemental analysis on recall content provides some additional insight into the auditors' problem representations. Across all three experience levels, more experienced auditors recalled a smaller proportion of general client statements and a larger number of statements related to financial information. The increase in financial recall measured as a proportion, however, was significant only when comparing juniors to the other two experience levels. One speculation about this finding might be that, with experience, auditors place more emphasis on analytical procedures in forming their problem representations. If this were the case, their recalls might include more references to financial information.

---

[17] When the REL statements were divided into subcategories and analyzed (see fn. 13), the managers and partners recalled a higher number and percentage of statements that related different types of information to each other. This result provides some additional limited support for the hypothesis related to knowledge of relationships.

Taken in total, the main finding of this study is that the planning problem representations of managers/partners may be different from those of seniors and juniors. Both experimental and anecdotal evidence indicates that senior auditors typically perform the planning function. In a background questionnaire, subjects in this study were asked, "In your firm, at what position or rank does an auditor typically assume the primary responsibility for planning the audit?" The responses were as follows:

Senior	80.9%
Manager	9.1%
Senior/Manager Team	7.2%
Other	2.8%

Therefore, the findings of this research may have important implications for practice. If differences in recall between managers/partners and seniors relate to "better" problem representations,[18] then more effective or efficient audits *may* be planned if managers and partners were actively involved in the process. However, further research will be needed to address this issue.

Recent criticisms of the auditing expertise literature (Ashton 1991; Bonner and Lewis 1990; Davis and Solomon 1989) have stressed the importance of identifying where knowledge differences exist and when such differences affect task performance. This study has identified some potential differences in auditor problem representations. However, it may be that the problem representations and planning knowledge of seniors are adequate for efficient, effective audit planning (Hogarth 1991). The question of how the information in the problem representations is processed and used must be addressed before these issues are resolved.

Appendix
Coding Examples

Recall Coded by Character

Basic facts (FAC) are direct statements or simple paraphrasing of information specifically stated in the case. Examples are:

"manufactures microcomputers"
"250,000 in cash"
"Director of Marketing worked for a pharmaceutical company before Fletcher"
"No exceptions on internal control questionnaire"
"No problems with previous auditors"

Relationships (REL) are statements that relate information in the case; these may include summarizing two or more items of information of the same type or relating two or more items of information of a different type. Examples are:

"Bad debts are up"
"The company has grown each year for the past three years"
"A/R aging has more >90 days since 12/86"

---

[18] Conceptually, there may be a "best" (or set of best) problem representation for a given audit situation. Such a representation would include an understanding of all the relationships and abstractions necessary to plan an effective and efficient audit.

"All VPs are over 35"
"A/R up in spite of lower sales"
"Accrued compensation appears to be increasing even though earnings are down"
"Experiencing decreasing sales with no decrease in COGS"

Abstractions (ABS) are statements that include information not presented in the case; these may involve adding an evaluative component, such as an adjective or adverb, to the case information or going beyond the case information by suggesting explanations of the case material, asking questions about missing information, or making statements about auditing the company. Examples are:

"Has strong internal controls"
"Qualified knowledgeable people at the management positions"
"involved in a volatile industry"
"credit problems?"
"market share small & perhaps disproportionate to profitable sales potential"
"Obsolete technology a problem for realizability"
"Internal controls . . . on the surface look like they can be relied on"

Recall Coded by Content

Statements about any of the basic client and industry information in the case are categorized as general client (GC). Examples are:

"manufactures microcomputers"
"home office in Dallas"
"Audit to be completed by Feb. 15"
"Controller has 5 years of public accounting experience"
"President owns 20% of stock"

Statements about the information provided in the comparative financial statements or accounts receivable agings and distributions are categorized as financial (FS). Examples are:

"A/R ≈ 45,000,000 at 9-30-87"
"Bad debts are up"
"$341,000,000 balance sheet"
"Accounts Rec agings show the largest amount in 50–100 thousand category"
"Slight erosion in agings from prior year"

Statements about the general system design and the internal control procedures are categorized as internal control (IC). Examples are:

"Detail revenue cycle controls over transaction processing exist"
"credit terms are 30 days"
"strong control system"
"cash receipts are deposited intact"
"Sales order confirmation is sent to customer"

Statements about information not in the case are categorized as other (OT). Examples are:

"Inventory at risk of obsolescence"
"Internal audit staff needs better direction and changed emphasis"
"Fletcher . . . appears aggressive"
"Collectibility of Receivables—major problem & concern"
"No mention of mgmt integrity"

References

Ashton, A. H. 1991. Experience and error frequency knowledge as potential determinants of audit expertise. *The Accounting Review* 66 (April): 218–39.

——, D. N. Kleinmuntz, J. B. Sullivan, and L. A. Tomassini. 1988. Audit decision making. In *Research Opportunities in Auditing: The Second Decade*, edited by A. R. Abdel-khalik and I. Solomon. Sarasota, FL: AAA.

Bedard, J. C., and S. F. Biggs. 1991. Pattern recognition, hypotheses generation, and auditor performance in an analytical task. *The Accounting Review* 66 (July): 622–42.

Biggs, S. F., T. J. Mock, and P. R. Watkins. 1988. Auditors' use of analytical review in audit program design. *The Accounting Review* 63 (January): 148–61.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Bonner, S. E., and B. L. Lewis. 1990. Determinants of auditor expertise. *Journal of Accounting Research* 28 (Supplement): 1–20.

——, and N. Pennington. 1991. Cognitive processes and knowledge as determinants of auditor expertise. *Journal of Accounting Literature* 10: 1–50.

Bouwman, M. J. 1982. The use of accounting information: Expert versus novice behavior. In *Decision Making: An Interdisciplinary Inquiry*, edited by G. R. Ungson and D. N. Braunstein. Boston, MA: Kent.

Bransford, J. D., and M. K. Johnson. 1972. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior* 11 (December): 717–26.

Brewer, W. F., and G. V. Nakamura. 1984. The nature and functions of schemas. In *Handbook of Social Cognition*, vol. 1, edited by R. S. Wyer and T. K. Srull. Hillsdale, NJ: Erlbaum.

——, and J. C. Treyens. 1981. Role of schemata in memory for places. *Cognitive Psychology* 13 (April): 207–30.

Carroll, J. M., J. C. Thomas, and A. Malhotra. 1980. Presentation and representation in design problem-solving. *British Journal of Psychology* 71 (February): 143–53.

Chi, M. T. H., P. Feltovich, and R. Glaser. 1981. Categorization and representation of physics problems by experts and novices. *Cognitive Science* 5 (April): 121–52.

——, R. Glaser, and E. Rees. 1982. Expertise in problem solving. In *Advances in the Psychology of Human Intelligence*, edited by R. J. Sternberg. Hillsdale, NJ: Erlbaum.

Chiesi, H. L., G. J. Spilich, and J. F. Voss. 1979. Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior* 18 (June): 257–73.

Choo, F., and K. T. Trotman. 1991. The relationship between knowledge structure and judgments for experienced and inexperienced auditors. *The Accounting Reveiw* 66 (July): 464–85.

Cushing, B. E., and J. K. Loebbecke. 1986. *Studies in Accounting Research No. 26: Comparison of Audit Methodologies of Large Accounting Firms.* Sarasota, FL: AAA.

Davis, J. S., and I. Solomon. 1989. Experience, expertise, and expert-performance research in public accounting. *Journal of Accounting Literature* 8: 150–64.

Dellarosa, D., and L. E. Bourne, Jr. 1984. Decisions and memory: Differential retrievability of consistent and contradictory evidence. *Journal of Verbal Learning and Verbal Behavior* 23 (December): 669–82.

Felix, W. L., and W. R. Kinney, Jr. 1982. Research in the auditor's opinion formulation process: State of the art. *The Accounting Review* 57 (April): 245–71.

Frederick, D. M. 1991. Auditors' representation and retrieval of internal control knowledge. *The Accounting Review* 66 (April): 240–58.

——, and R. Libby. 1986. Expertise and auditors' judgments of conjunctive events. *Journal of Accounting Research* 24 (Autumn): 270–90.

Glaser, R. 1984. Education and thinking: The role of knowledge. *American Psychologist* 39 (February): 93–104.

——, and M. T. H. Chi. 1988. Overview. In *The Nature of Expertise*, edited by M. T. H. Chi et al. Hillsdale, NJ: Erlbaum.

Guilford, J. P., and B. Fruchter. 1978. *Fundamental Statistics in Psychology and Education.* New York: McGraw-Hill.

Hogarth, R. M. 1991. A perspective on cognitive research in accounting. *The Accounting Review* 66 (April): 277–90.

Johnson, P. E., K. Jamal, and R. G. Berryman. 1991. Effects of framing on auditor decisions. *Organizational Behavior and Human Decision Processes* 50 (October): 75–105.

Kassirer, J. P., and R. I. Kopelman. 1987. The critical role of context in the diagnostic process. *Hospital Practice* 22 (August): 67–76.

——, ——. 1989. Cognitive errors in diagnosis: Instantiation, classification, and consequences. *The American Journal of Medicine* 86 (April): 433–41.

Kinney, W. R., Jr. 1986. Audit technology and preferences for auditing standards. *Journal of Accounting & Economics* 8 (March): 73–89.

Lesgold, A., H. Rubinson, P. Feltovich, R. Glaser, D. Klopfer, and Y. Wang. 1988. Expertise in a complex skill: Diagnosing X-ray pictures. In *The Nature of Expertise*, edited by M. T. H. Chi et al. Hillsdale, NJ: Erlbaum.

Libby, R. 1985. Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research* 23 (Autumn): 648–67.

———, and D. M. Frederick. 1990. Experience and the ability to explain audit findings. *Journal of Accounting Research* 28 (Autumn): 348–67.

Mandler, J. M. 1984. *Stories, Scripts, and Scenes: Aspects of Schema Theory.* Hillsdale, NJ: Erlbaum.

———, and N. S. Johnson. 1977. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology* 9 (January): 111–51.

McKeithen, K. B., J. S., Reitman, H. H. Rueter, and S. C. Hirtle. 1981. Knowledge organization and skill differences in computer programmers. *Cognitive Psychology* 13 (July): 307–25.

Moeckel, C. 1990. The effect of experience on auditors' memory errors. *Journal of Accounting Research* 28 (Autumn): 368–87.

Newell, A., and H. Simon. 1972. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Norman, G. R., L. R. Brooks, and S. W. Allen. 1989. Recall by expert medical practitioners and novices as a record of processing attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15 (November): 1166–74.

Pennington, N. 1987. Stimulus structure and mental representations in expert comprehension of computer programs. *Copgnitive Psychology* 19 (July): 295–341.

Rumelhart, D. E., 1977. Understanding and summarizing brief stories. In *Basic Processes in Reading: Perception and Comprehension*, edited by D. LaBerge and S. J. Samuel. Hillsdale, NJ: Erlbaum.

———. 1984. Schemata and the cognitive system. In *Handbook of Social Cognition*, vol. 1, edited by R. S. Wyer and T. K. Srull. Hillsdale, NJ: Erlbaum.

Sandlin, P. 1987. *The Effect of Causal Background on Auditors' Analytical Review Judgments.* Unpublished dissertation, University of Texas, Austin.

Tubbs, R. M. 1992. The effect of experience on the auditor's organization and amount of knowledge. *The Accounting Review* 67 (October): 783–801.

Weber, R. 1980. Some characteristics of the free recall of computer controls by EDP auditors. *Journal of Accounting Research* 18 (Spring): 214–41.

Winer, B. J. 1971. *Statistical Principles in Experimental Design*, 2d ed. New York: McGraw-Hill.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 323–345

Strategic Sampling, Physical Units Sampling, and Dollar Units Sampling

Stephen C. Hansen
Stanford University

SYNOPSIS AND INTRODUCTION: One of the most common decisions facing an internal auditor is choosing which line items to investigate. An extensive literature (Dworin and Grimlund 1984; Leslie et al. 1980; Menzefricke 1984; Teitlebaum and Robinson 1975) deals with the statistical and decision-theoretic aspects of his choice. This paper expands on previous work by adding a strategic source of errors: dishonest employees. It addresses the question of how the presence of strategic errors affects the relationship between the auditor's testing strategy and item value. I show that incorporating strategic errors can lead to audit strategies similar to Physical Units and Dollar Units Sampling.

I highlight the assumptions driving the results by contrasting a firm's (or internal auditor's) use of an optional test in four stylized models of accounts receivable.

The first model examines the firm's behavior when faced with nonstrategic (statistical) billing errors. In this model the accounting system generates random errors that result in over- or underbilling customers. The firm can use a costly, imperfect test to remove errors before the bills are sent out. In this nonstrategic model the firm randomizes and tests an item if and only if the benefit is greater than the cost. Because the amount of billing error is unrelated to the item value, there is no clear link between the firm's testing decision and the value of the line item.

The second billing model adds the possible existence of dishonest employees who can steal from line items. A dishonest employee makes two decisions. He decides whether to steal from the line item, and, if he steals, he chooses the amount of the theft.

A dishonest employee would steal the entire item if he were certain

I would like to thank Steven Huddart, Bjørn Jørgensen, Susanne Lohmann, Nahum Melumad, Paul Newman, Bharat Sarath, Toshiyuki Shibano, Konduru Sivaramakrishnan, participants in a seminar at the University of Texas at Austin, two anonymous referees, and an associate editor for insightful comments, and the Accounting Research Center at Northwestern University for financial support.

Submitted February 1991.
Accepted August 1992.

that the firm would never test that item. The dishonest employee's behavior forces the firm to consider the value of the item in determining the region of untested items. Specifically, low value items are never tested. As in many strategic models, the interaction with dishonest employees may lead to randomization. In particular, the randomized testing strategy can look like Stratified Physical Units Attributes Sampling (Leslie et al. 1980). The firm sorts items into different groups and each item in a group has the same probability of being tested.

The third model contains only the statistical errors of incorrectly adding or deleting a sales discount, a percentage of the item value. Since the testing gain is directly related to the value of the line item, the firm's strategy depends on an item's value. The firm always tests high value items, and never tests low value items.

The fourth model adds potentially dishonest employees who can provide unearned sales discounts to their confederates. In this model the firm stratifies items into *three* groups. It never investigates small items, always investigates large items, and randomizes over intermediate value items with probabilities roughly proportionate to the value of the item. This procedure is similar to a common audit procedure, Dollar Unit Cell Width Sampling (Leslie et al. 1980).

Key Words: *Strategic auditing, Game theory, Dollar unit sampling, Accounts receivable.*

Data Availability: *Complete proofs of all theorems are available from the author.*

A NUMBER of papers (Anderson and Young 1988, Fellingham et al. 1989; Fellingham and Newman 1985; Shibano 1990) introduce strategic play into the auditing framework. The common approach is to focus on the manager's effect on the error rates of the line items. These models assume that either the auditor investigates one account (Fellingham and Newman 1985; Shibano 1990) or that the auditor's and manager's payoffs are determined by the error rate in the population (Fellingham et al. 1989; Patterson 1990).[1]

While these models offer insight into the effects of a strategic player, they operate at a very general level. Because all items have identical characteristics, they cannot explore the relationship between the strategies and "micro" characteristics of the items, such as the value and the distribution of the error amount. In order to make explicit statements about the link between the testing strategy and item characteristics, I impose more structure on the model. I examine a smaller slice of the firm's activities and investigate a firm's or internal auditor's use of an optional test in stylized models of accounts receivable. I examine the firm's use of a control when there are multiple items (bills) which can vary in value and other attributes.

The presentation is organized in four sections. In the first section, the nonstrategic, billing errors model is introduced. In the second section the strategic billing model is analyzed. Section III investigates the strategic and nonstrategic sales discount models. The final section discusses the results.

---

[1] An exception is the paper by Newman et al. (1991).

I. The Nonstrategic Model with Statistical Billing Errors

The first model examines a firm's optimal testing strategy when there are only statistical errors and there is no clear connection between the item value and the net gain from testing.

A risk neutral firm wishes to prevent billing errors for its accounts receivable. The firm has N line items in accounts receivable with stated dollar values $B_1 > B_2 > \ldots > B_N > 0$. In order to simplify the analysis, I assume that these values reflect the true amount of a customer's bill and would be paid if sent to the customer.[2] There is a probability that errors occur in processing the bills and sending them to the customers. A crucial feature is that the errors are nonstrategic: they do not change as the firm modifies its strategy.

To motivate further assumptions about the errors in line items, I use the empirical evidence found in Johnson et al. (1981). Johnson examined the probabilities and error distributions for the line items of accounts receivable and found three empirical regularities: (1) the probability of an error is increasing in the value of the line item; (2) given that an error has occurred, the mean of the error amount distribution is not increasing in the value of the line item; and (3) given that an error has occurred, the variance of the error amount distribution is increasing in the value of the line item.

In my model, the probability of an error is $a_i > 0$ for each line item i $(i = 1, \ldots, N)$. In keeping with the empirical evidence, the probability an error exists is (weakly) increasing in the value, $1 > a_1 \geq \ldots \geq a_N > 0$.

The probability of an error and the amount of the error are determined independently. Given that an error has occurred in processing line item i, the customer will receive a bill of value $B_i + \tilde{e}_i$. The error term $\tilde{e}_i$ is drawn from a probability density function $f_i(e)$ with cumulative distribution function $F_i(e)$ and support $(\underline{B}, \bar{B})$. The error amount distribution allows for the possibilities of over- or underbilling of customers and assumes that there may be different distributions for different line items. The distribution is bounded, since the control system will detect immediately extraordinarily large or small errors. While I allow different potential error amount distributions for different items, I assume that all the error amount distributions have a common mean of $\bar{e} = 0$, and that the variance is increasing in the value of the item.

Returned bills involve reprocessing costs, and undetected underbilling (overbilling) involves losses (gains) equal to the value of the error. Since the company is collecting revenues from customers, undetected errors affect the firm's profits even if the company never discovers the errors exist. Customers have a higher probability of returning overbilling, *hi*, than underbillings, *lo*, since customers may notice underbillings and decide to pay the lower bill, that is, $hi > lo > 0$.[3] Because customers are "virtuous" in returning underbilled items, the rebilling loss is lower for returned underbilled items, *rl*, than overbilled items, *rh*, that is, $rh > rl$.

The following list summarizes, the potential losses and gains to the firm:

If $\tilde{e}_i > 0$, then	with probability hi	the firm loses rh,
	with probability $(1 - hi)$	the firm gains $\tilde{e}_i$.

---

[2] I conjecture that the results would not change if the stated dollar balances could be in error. I can reinterpret the initial dollar balances as the best estimate of the value of the line item, and a similar analysis should go through.

[3] A consequence of the customer's bias is that the distribution of undetected errors will be skewed toward underbillings, even though the mean of the amount of error distribution is zero.

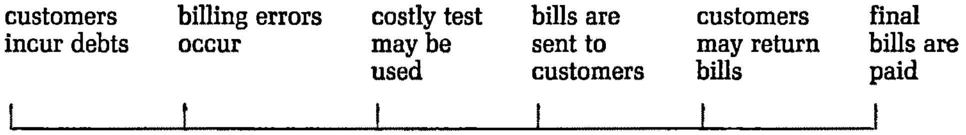
If $\tilde{e}_i<0$, then with probability lo the firm loses rl,
with probability $(1-lo)$ the firm loses $-\tilde{e}_i$.

My model is concerned with the firm's optimal response to the possibility of errors. An implicit assumption is that all supervisors have incentives aligned with those of the firm. The internal auditor designs the control system, and the manager uses the control system in a fashion that maximizes the value of the firm.

The control system of the firm contains multiple checks on the accounts receivable process.[4] The firm has implemented several of these controls and is contemplating adding another. The firm has the possibility of using an additional, costly, internal control to remove errors before bills are sent to the customers. This control is a prevention control. It decreases the probability of incorrect bills being sent to customers. A key feature of this control is that it is not necessarily used for all transactions. For example, this control could consist of the manual checking of invoices, or the dual processing of transactions.

The control costs C_i when it is applied to item i. Larger line items (orders) typically have more components on their bills; therefore the larger line items will have higher testing costs, $C_1\geq\ldots\geq C_N>0$. Given that the control is used, the probability that an error is detected for line item i is p_i, where $1>p_i>0$.[5]

The following timeline captures the sequence of events in the model.



Before proceeding further, it is necessary to write down several definitions that simplify the notation used later. The first two definitions are the *ex ante* expected amounts of overbilling (underbilling) given that an overbilling (underbilling) error has occurred,

$$O_i\equiv\frac{1}{[1-F_i(0)]}\int_0^{\bar{B}} ef_i(e)de, \qquad U_i\equiv\frac{1}{F_i(0)}\left|\int_{\underline{B}}^{0} ef_i(e)de\right|.$$

The third definition is the expected loss from an undetected billing error,

$$\ell_i\equiv[1-F_i(0)][(hi)rh-(1-hi)O_i]+F_i(0)[(lo)rl+(1-lo)U_i].$$

In order to motivate the firm to use the control, I assume that the firm has an expected loss from each undetected error, that is, $\ell_i>0$ for all i.[6]

---

[4] Wallace (1984, chap. 5) describes the basic types of controls as segregation of incompatible functions, creating an audit trail by the correct use of forms and documents, and having an effective environment for control.

[5] I assume that there is no cost for reprocessing bills with errors detected by the optional control. A faulty bill detected by the customers involves much higher costs than a faulty bill detected in the accounting process. The firm loses goodwill when customers must correct the errors. Setting the cost for reprocessing internally detected errors equal to zero simplifies the model while preserving the hierarchy between the two types of detected errors.

[6] This assumption is necessary since the firm gains if it successfully overbills customers. One possibility is that the gains could outweigh the losses and the firm would desire billing errors.

The next step is to write down the firm's minimization problem. The firm chooses how to use the additional preventative control. In my model, the firm minimizes the sum of expected losses and the costs of testing. If $T_i(i=1,\ldots,N)$ is the probability that the firm tests item i, then the firm solves,

$$\min_{T_1,\ldots,T_N} \sum_{i=1}^{n} \{T_i[a_i(1-p_i)\ell_i+C_i]+(1-T_i)a_i\ell_i\},$$

subject to:

$$1 \geq T_i \geq 0 \qquad \text{for } i=1,\ldots,N.$$

The firm's objective function contains three elements. The first, $T_i a_i(1-p_i)\ell_i$, reflects the expected loss when the firm tests, a statistical error occurs, and the firm fails to detect the error. The second element, $T_i C_i$, captures the testing cost paid by the firm if it tests the item. The final element, $(1-T_i)a_i\ell_i$, specifies that the firm also incurs a loss if it does not test and a statistical error occurs.

The solution to the firm's program is straightforward. Due to the risk neutrality (linearity) of the firm's preferences, the firm always tests or never tests any specific line item. It tests an item if the decrease in expected losses outweighs the cost of testing. It tests if the net gain from testing is greater than zero. The optimal rule is to test an item if and only if,

$$a_i p_i\{[1-F_i(0)][(hi)rh-(1-hi)O_i]+F_i(0)[(lo)rl+(1-lo)U_i]\}-C_i \geq 0. \tag{1}$$

Equation (1) provides insight into the conditions leading to testing. The firm will test an item if the probability of an error is high, the probability of detecting an error is high, the cost of rebilling is high, or the cost of testing is low. These are all inherent properties of the line item (customer) and the firm's accounting system.

An important question is how the firm's testing strategy is related to the value of the item. The following two examples indicate that the relationship between the strategy and value is ambiguous.

Example 1. Assume the expected loss due to an undetected error, ℓ_i, and the detection probability, p_i, are constant for all items, and that the probability of a statistical error, a_i, rises faster than the cost of testing, C_i. In this situation, the auditor's optimal strategy will be to always test the items whose values lie above a threshold. Figure 1 illustrates this strategy.

Example 2. Assume the expected loss due to an undetected error, ℓ_i, and the detection probability, p_i, are constant for all items, and that the probability of a statistical error, a_i, rises slower than the cost of testing, C_i. In this situation, the auditor's optimal strategy will be to always test the items whose values lie below a threshold. Figure 2 presents this strategy.

Many other examples can be generated with many other patterns. However, these two examples illustrate that the connection between item value and testing strategy depends on the parameters of the model.

II. The Strategic Model with Statistical Billing Errors and Employee Irregularities

The second model maintains the assumption that there is no clear connection between the net testing gain for statistical errors and item value. However, the model

Figure 1
The Statistical Billing Errors Model, Example 1

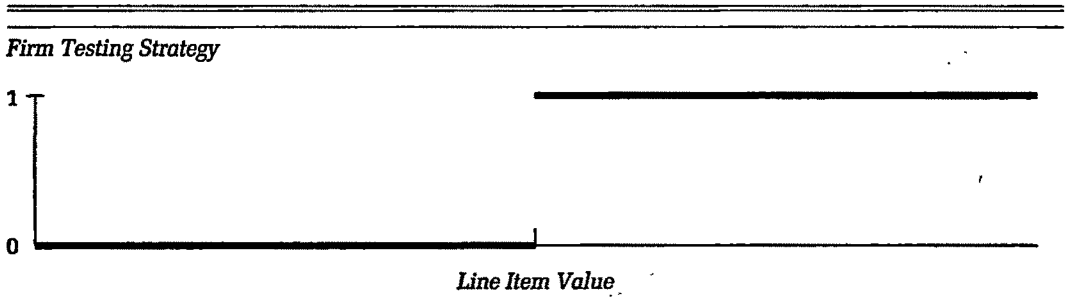
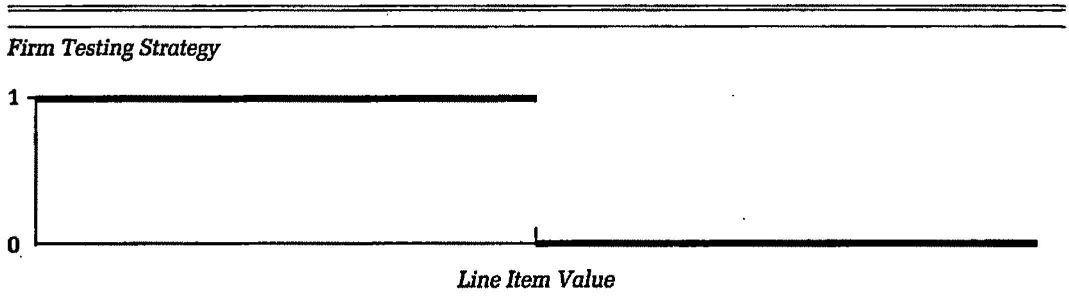


Figure 2
The Statistical Billing Errors Model, Example 2



modifies the error structure so that now the discrepancies can have two sources. Now items can be missing due to either statistical errors or employee irregularities. In other words, there are now nonstrategic (statistical) and strategic (employee) sources of error.[7]

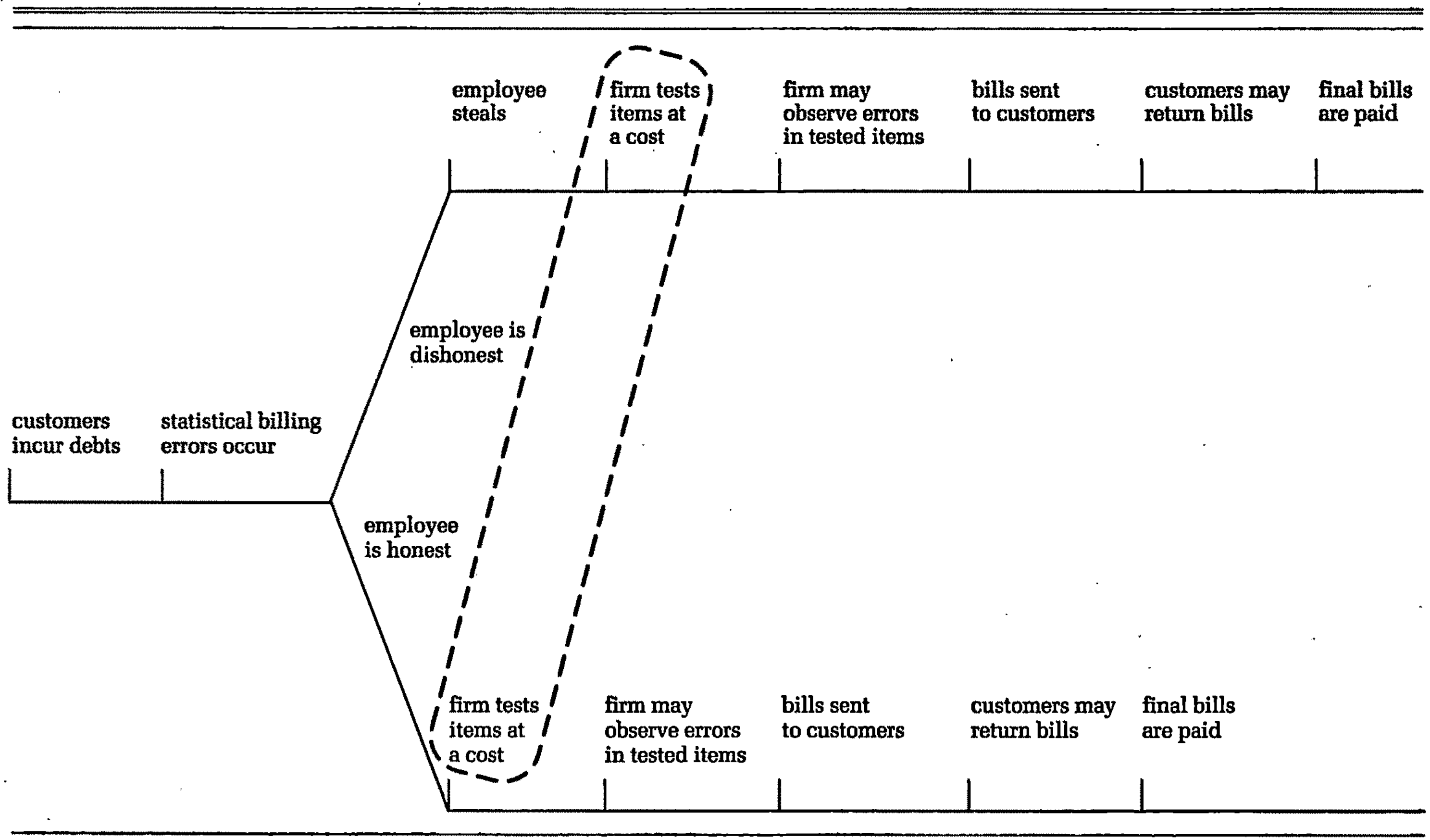
The crucial difference between this model and the previous one is the addition of a strategic interaction between the firm and dishonest employees. Both the firm's testing strategy and the employees' theft strategies (the distribution of strategic errors) are the equilibrium results of a game.

The sequence of this game is straightforward (see fig. 3). Nature initially draws a set of statistical errors for each of the N line items. The firm does not observe these errors. There are N employees, one for each line item, and nature then determines whether each employee is dishonest. The probability of a dishonest employee is $1 \geq \delta_i > 0$. The firm cannot observe whether employees are honest or dishonest. The dishonest employees' move next (if they exist). Each dishonest employee chooses and implements a theft strategy. After these thefts have occurred, the firm chooses a testing strategy and pays the costs of testing. The firm then observes the outcome of the noisy detection

---

[7] A model with only strategic irregularities has the same qualitative features as the model with both statistical errors and strategic irregularities.

Figure 3
The Timeline for One Line Item


customers incur debts
statistical billing errors occur
employee is dishonest
employee is honest
employee steals
firm tests items at a cost
firm may observe errors in tested items
bills sent to customers
customers may return bills
final bills are paid
firm tests items at a cost
firm may observe errors in tested items
bills sent to customers
customers may return bills
final bills are paid

technology and corrects the observed errors. Bills are then sent to the customers who may return them for reprocessing. Finally, all bills are paid.

An important point is that the probability of dishonest employees is not one. Most employees could be honest, but it is the possibility that some may be dishonest which leads to the firm's strategic behavior. The firm may screen employees for honesty, but I assume that no screen is perfect and some dishonest employees may have slipped through.

A dishonest employee has a fairly straightforward theft technique. Each dishonest employee has made a pact with the individual whose bill he processes. If the dishonest employee decides to steal, he reduces the prices and/or quantities charged to his customer. The customer reciprocates by giving the dishonest employee a percentage, α, of the amount reduced.[8,9]

The dishonest employee has two choices. He chooses the probability that he reduces line item i, S_i, and, if he steals, he decides the size of the theft, V_i.[10] The dishonest employee cannot reduce the line item by more than the value of the line item, $B_i \geq V_i$.[11] If he steals amount V_i and the firm tests that account, then the probability that the theft is detected is $d(V_i)$. The detection probability increases with the amount stolen, $d'(\bullet)>0$. If the theft is detected, the employee pays a liability, $L(V_i)$, that increases in the size of the amount stolen, that is $L'(\bullet)>0$. The employee's liability can include firing and civil penalties, as well as the psychic costs of observing the punishment of his confederate. If the dishonest employee does not steal from the account, then the control will not mistakenly detect a theft or mistakenly generate a liability, $d(0)=L(0)=0$.

Combining both statistical errors and a dishonest employee requires several assumptions about the interaction between the two. The presence of a poor control system induces greater employee dishonesty.[12] Therefore, the probability of a dishonest employee is assumed to be increasing in the level of statistical error, $\delta_i'(a_i)>0$. Although the probability of a dishonest employee and the probability of a statistical error are related, the draws from the two distributions are assumed to be independent. The distinction between the probability of detecting a statistical error and detecting an irregularity reflects the idea that the dishonest employee will attempt to cover up his theft. If there is only a statistical error in an account, the probability of detection remains p_i. If there is a strategic error in an account, I assume that the probability of detection is $d(V_i)$.

---

[8] Sawyer (1988, 1016–17) lists 40 common forms of fraud. Several elements of the list are:

24. Paying false invoices, either self-prepared or obtained through collusion with suppliers.
25. Increasing amounts of suppliers' invoices through collusion.
40. Permitting special prices or privileges to customers, or granting business to favored suppliers, for "kickbacks."

[9] After the employee adjusts the bill for the customer, the customer may wish to renege on the bargain. One possible reason the customer may not do so is that the arrangement may be long term. The customer may follow through because he wishes to obtain bill reductions in the future.

[10] This dishonest employee incurs no costs in carrying out his theft.

[11] This restriction bounds the potential losses to the firm and is needed in the proof of Theorem 1.

[12] Hollinger and Clark (1983, 117) state:

> The good news is that employee theft does seem to be susceptible to control efforts. In both retail firms and hospitals, various controls were related consistently to lowered rates of theft by employees. . . . Our data also indicate, however, that the impact of organizational controls is neither uniform nor very strong.

The dishonest employee does not know if a statistical error is present when he reduces the customer's bill. However, the customer is a confederate of the dishonest employee and will not wish to draw attention to any reduced bill. Therefore, the customer will not return any bill which the dishonest employee has modified.[13]

In this model the firm cannot observe the presence of a dishonest employee or his theft strategy when making its decisions. Similarly, a dishonest employee cannot observe the firm's testing strategy when deciding how to steal. However, this setting is not a simultaneous move game because of a fundamental information asymmetry. The employee knows that he is dishonest, but the firm does not know whether it faces a dishonest employee. This asymmetry means the model must be analyzed as a sequential move game.

The use of a mathematical game requires an equilibrium concept. The one chosen is subgame perfect Nash. The firm chooses its strategy taking as given the dishonest employee's strategy. Similarly, the dishonest employee chooses his strategy taking as given the firm's strategy. In equilibrium, neither the firm nor the dishonest employee wishes to change their strategy.

A dishonest employee is motivated solely by economic, not moral, considerations and wishes to maximize the net value of his thefts.[14] The dishonest employee for line item i chooses the probability S_i and the size of the theft V_i to solve,

$$\max_{S_i,\,V_i} \quad S_i\{\alpha V_i[1-T_i d(V_i)]-T_i d(V_i)L(V_i)\},$$

subject to:

$$B_i \geq V_i \geq 0,$$
$$1 \geq S_i \geq 0.$$

The employee's objective function is straightforward. The first term, $\alpha V_i[1-T_i d(V_i)]$, indicates that the employee receives his share of the overbilling if he steals and is not detected. The second term, $T_i d(V_i)L(V_i)$, captures the employee's liability if he steals and is detected.

Implicit in the dishonest employee's payoff is the assumption that if the firm detects an error, it can determine the source of the error. If an employee steals from an item and the error is detected, then the error is correctly classified as theft. This assumption is used for convenience. Adding in a probability that an employee irregularity is incorrectly classified as a statistical error increases the number of terms in the objective functions, but does not change the qualitative features of the results.[15]

---

[13] I assume that overbillings are never large enough that the customer will defect and turn in the dishonest employee. The previous assumption that overbillings are bounded, that extremely large overbillings are always detected, offers one justification for this assumption.

[14] Bologna and Lindquist (1987, 32), state:

> As a general rule, [criminal] motives can be separated into four major categories: (1) psychotic, (2) economic, (3) egocentric, and (4) ideological. Psychotic motivation can impair a successful criminal prosecution (e.g., when it calls for an insanity plea). Of the other three motivations, economic motivation is the most common.

[15] If the firm must incur an additional cost to distinguish errors from irregularities, then this cost could lead to more theft. The firm would never wish to differentiate between small thefts and statistical errors. Therefore, the dishonest employees always would steal small amounts just at the cost threshold. However, my model still applies to thefts above this threshold.

Now that the dishonest employee's maximization problem has been stated, the firm's minimization problem will be presented. As in the previous model, the firm minimizes its expected losses and testing costs.[16] It chooses a set of testing probabilities $T_i(i=1,\ldots,n)$, that solve,

$$\begin{aligned}\min_{T_1,\ldots,T_N} \quad & \sum_{i=1}^{n}\Big\{T_i\Big[\delta_i S_i[1-d(V_i)]\Big(V_i+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\Big)\\ & +\delta_i(1-S_i)a_i(1-p_i)\ell_i+(1-\delta_i)a_i(1-p_i)\ell_i+C_i\Big]\\ & +(1-T_i)\Big[\delta_i S_i\Big(V_i+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\Big)\\ & +\delta_i(1-S_i)a_i\ell_i+(1-\delta_i)a_i\ell_i\Big]\Big\}\end{aligned}$$

subject to:

$$1\geq T_i\geq 0 \qquad \text{for } i=1,\ldots,N.$$

The firm's objective function contains several parts. The first part, the T_i term, contains four elements. The first element captures the case where the firm tests, the employee is dishonest, the employee decides to steal, and the firm fails to detect the theft. The probability that all four occur is $T_i\delta_i S_i[1-d(V_i)]$. In this case, the amount of the loss is the value of theft, V_i, plus the additional expected losses of a statistical error, $a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}$. The expected loss due to a statistical error differs from the loss in the prior model because of the assumption that when theft occurs, the customer is a confederate of the employee. The confederate will never return a bill with a statistical error, and therefore the firm never incurs any rebilling costs. The second and third elements of the T_i term indicate that the firm also loses if it tests, the dishonest employee does not steal, and a statistical error occurs, $T_i\delta_i(1-S_i)a_i(1-p_i)\ell_i$, or if it tests, the employee was honest, and a statistical error occurs, $T_i(1-\delta_i)a_i(1-p_i)\ell_i$. The final element, T_iC_i, specifies that the firm pays the cost of testing only when it tests an item.

The $(1-T_i)$ term has two modifications relative to the T_i term. Since the firm does not test the items, both the probability of failing to detect an error and the cost of testing are absent.

Now that both the firm and the dishonest employee's optimizaton problems have been presented, I will analyze the equilibrium. In order to simplify the analysis, I assume that the expected net gain from theft,

$$\alpha V_i-T_i d(V_i)[\alpha V_i+L(V_i)],$$

is strictly concave for all $T_i>0$ and has a unique maximum at V^{*T}, where $B_i>V^{*T}>0$ for $T_i>0$. This assumption says that the dishonest employee has a finite satiation point in theft for each positive testing probability. The satiation amount reflects the optimal tradeoff between the return from a successful theft and the potential loss if caught.

---

[16] The firm's objective function implies that the firm can recover the amount of any detected theft. Making the alternate assumption that the dishonest employee and his confederate immediately "consume" some or all of the value of the theft does not change the nature of the results. However, if there is immediate consumption of the thefts, the firm will increase the level of testing in order to prevent a greater amount of theft than in the full recovery case.

The characteristics of the equilibrium depend on the parameters of the model. The following theorem presents the most interesting case.[17]

Theorem 1: If (1) the firm faces both strategic and nonstrategic errors, (2) the decline in the firm's loss when it detects a strategic irregularity is larger than the reduction when it detects a statistical error,[18,19]

$$d(V^{RS})\left(V^{RS}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\right)>a_ip_i\ell_i, \tag{2}$$

(3) the dishonest employee would not steal if the firm always tested the item,

$$\alpha V^{*1}-d(V^{*1})[\alpha V^{*1}+L(V^{*1})]<0,$$

and (4) $C^{T|S}>C^{RS}$,[20] then the equilibrium firm testing probabilities, employee theft probabilities, and theft amounts can be partitioned into the following regions:

1. If $C^{AT}>C_i$, then the firm always tests the line item, and the employee never steals.
2. If $C^{RS}>C_i>C^{AT}$, then firm tests the item with a constant probability T^{RS}, a dishonest employee steals with probability S_i, and he steals the constant amount V^{RS}. These values are the solution to the equations,

$$T^{RS}=\frac{\alpha}{d'(V^{RS})[\alpha V^{RS}+L(V^{RS})]+d(V^{RS})[\alpha+L'(V^{RS})]},$$

$$S_i=\frac{C_i-a_ip_i\ell_i}{\left[\delta_i d(V^{RS})\left(V^{RS}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\right)-\delta_i a_ip_i\ell_i\right]},$$

---

[17] The remaining cases of the theorem are provided in appendix A. The changes from Theorem 1 are relatively minor. Both cases 2 and 3 lack Region (2) of the theorem. Case 3 has a region where the firm always tests and the dishonest employees always steal.

[18] If equation (2) does not hold, thefts reduce the firm's expected losses and are a net benefit. The employee steals smaller amounts than the statistical errors generated by the billing process, and the firm would prefer thefts by the dishonest employee.

[19] The term

$$d(V_i)\left(V^{RS}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\right)$$

represents the firm's reduced losses due to detecting an irregularity, while the term $a_ip_i\ell_i$ captures the firm's reduction in losses due to detecting a statistical error in the nonstrategic model.

[20] The cutoff C^{NT} (Never Test) is the cost where the firm is indifferent between testing and not testing given that the employee will always steal the total amount B_i. If the cost is greater than C^{NT}, then the firm will not test that line item.

The cutoff $C^{T|S}$ (Test given Stealing) is the cost where the firm is indifferent between testing and not testing given that the employee will always steal the amount V^{*1}. If the cost is below $C^{T|S}$ and the employee steals from this item, then the firm will always test the line item. The difference between $C^{T|S}$ and C^{NT} is due to the discontinuity in the employee's payoff when the firm stops investigating the line item. If the firm always tests a line item and a dishonest employee always steals, then a dishonest employee would steal the interior amount V^{*1}. If the firm never tests a line item and the dishonest employee always steals, then a dishonest employee would always steal the maximum possible, B_i.

The cutoff C^{AT} (Always Test) is the cost where the firm is indifferent between testing and not testing assuming that the employee would never steal. If the cost is below C^{AT}, then the firm always tests the line item. The value of reducing the statistical error is sufficient to justify testing.

The final cutoff is derived from the employees' optimization problem. The cutoff C^{RS} (Randomized Stealing) is the cost where the employee is indifferent between stealing and not stealing from the line item. If the cost is less than C^{RS}, then the employee may randomize over stealing from the item.

and

$$\alpha V^{RS}-T^{RS}d(V^{RS})[\alpha V^{RS}+L(V^{RS})]=0.$$

3. If $C^{NT}>C_i>C^{RS}$, then the firm tests the item with probability:

$$T_i=\frac{\alpha}{d'(V^{*T})[\alpha V^{*T}+L(V^{*T})]+d(V^{*T})[\alpha+L'(V^{*T})]},$$

a dishonest employee always steals, and the size of the theft, V^{*T}, solves:

$$\delta_i d(V^{*T})\left(V^{*T}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\right)=C_i-(1-\delta_i)a_ip_i\ell_i.$$

4. If $C_i>C^{NT}$, then the firm never tests the line item, a dishonest employee always steals, and he steals the total line item.

The constants solve:

$$C^{NT}=\delta_i d(B_i)\left(B_i+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\right)+(1-\delta_i)a_ip_i\ell_i, \tag{3}$$

$$C^{T|S}=\delta_i d(V^{*1})\left(V^{*1}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\right)+(1-\delta_i)a_ip_i\ell_i,$$

$C^{AT}=a_ip_i\ell_i$, and

$$C^{RS}=\delta_i d(V^{RS})\left(V^{RS}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\right)+(1-\delta_i)a_ip_i\ell_i.$$

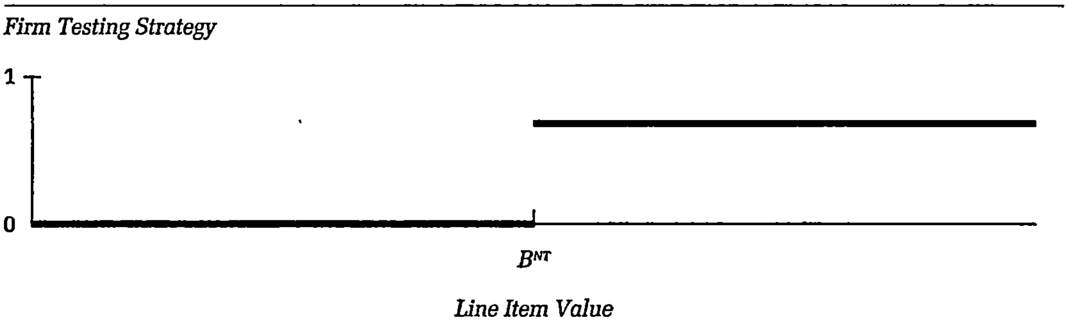
Proof: See appendix A.

The four regions specified in Theorem 1 represent four different equilibrium behaviors. In the first region, the firm's strategy of always testing the item deters the dishonest employees from any theft. In the second region the firm randomizes over items with a constant probability. The employee randomizes over stealing, and he steals a constant satiation amount. In the third region the firm randomizes over testing, the employee always steals, and he steals a satiation amount that varies with the item. The distinction between the second and third cases is that in the second the employee chooses a randomized theft probability and a constant theft amount, while in the third case the employee chooses a constant theft probability (one) and varying theft amount. In the final region, the firm never tests, the dishonest employee always steals, and he steals the entire item.

An important difference between the firm's behavior in the statistical and strategic billing models is that the firm may randomize in the strategic model. Randomized testing is a direct consequence of the strategic interaction between the firm and the dishonest employees.

The next step is to examine the relationship between the item value and the firm's testing strategy. Drawing general conclusions leads to the same difficulties as in the statistical billing errors model. Depending on the parameters, almost any pattern can emerge. However, there is one situation where item value and testing strategy are related.

Figure 4
The Strategic Billing Errors Model, Corollary 1



Corollary 1: If the firm groups items into strata with the same parameters, and if statistical errors alone would not lead to testing, then the firm uses a strategy similar to Stratified Physical Units Attributes Sampling (Leslie et al. 1980).[21] It never tests small items, and tests large items with a constant, common probability (which could be one).

Proof: Another way of looking at equation (3) is that there is a cutoff item value B^{NT}.[22] If the value of the item is below B^{NT}, then the firm will never test that item. These items are too small to warrant investigation even though this may lead to increased theft by the dishonest employee.

The firm's behavior is straightforward for items with values above the cutoff B^{NT}. It always tests these items with the same probability. The reason is that when employees steal from these items, they always steal the same satiation amount, an amount which is unrelated to the value of the item.[23]

Figure 4 illustrates the firm's testing strategy given in Corollary 1.

The connection between the firm's testing strategy and item value is due to two factors. First, all parameters but value have been held constant. This restriction makes it possible to isolate the effect of the dishonest employee. Second, potentially dishonest employees do care about item value. They will steal as much as possible if they know they will not be caught. The presence of greedy dishonest employees creates potential losses that depend on the item value. Since losses can depend on item value, the firm selects the set of tested items using their values.

In Corollary 1 the firm does not test items with values below a specific cutoff. The firm uses an endogenously derived materiality level. The intuition for the materiality cutoff is identical to that given by external auditors. Investigation of low value items is too expensive relative to the potential gains.

Now that the models with billing errors have been analyzed, I move on to investigating the models with sales discount errors.

---

[21] The technique stratifies items into different groups and each item in a group is tested with the same probability.

[22] The cutoff C^{NT} is the only one affected by a change in the value of the line item.

[23] This statement depends upon the assumption that all parameters except for item value are held constant.

III. The Nonstrategic and Strategic Models with Sales Discount Errors

The previous models captured a situation in which there was no straightforward connection between the net gain from testing for statistical error and the value of the line item. This section examines a setting in which a relationship to the line item value exists. The new models investigate sales discount errors. Sales discounts are a percentage reduction in the amount of a customer's bill if the customer makes an early payment. The errors in the accounting system consist of giving unwarranted discounts, or leaving off earned discounts. The size of the loss (gain) to the firm is the amount of the discount and is assumed to be ϵB_i, where $1>\epsilon>0$.

Since the procedure of recording a check's arrival is independent of the number on the check, many of the elements of the previous model now are constant across line items. In particular, the probability of an arithmetic error will be a, the distribution of the size of errors will be $F(\bullet)$, the probability of detecting an error will be p, and the cost of checking an item will be C.

One additional piece of notation simplifies the analysis. The expected loss given a statistical error has occurred will be written as,

$$\tau_i \equiv [1-F(0)][(hi)rh-(1-hi)\epsilon B_i]+F(0)[(lo)rl+(1-lo)\epsilon B_i].$$

Notice that even though many model parameters are constant across items, the loss still depends upon the item value through the term ϵB_i.

A more subtle modification is that in these models statistical errors and strategic irregularities cannot occur at the same time for the same item. The reason is that only one unwarranted discount can occur for each bill. Either the dishonest employee provided it, or the accounting system generated it. Unlike the billing errors model, if both an error and irregularity occur for the same item, the existence of a "kickback" means it is classified as an employee irregularity.

The Statistical Sales Discount Model

When the firm only faces statistical discount errors, it chooses testing probabilities, $T_i (i=1,\ldots,n)$, to solve:

$$\min_{T_1,\ldots,T_N} \quad \sum_{i=1}^{n} [T_i a(1-p)\tau_i+(1-T_i)a\tau_i+T_iC_i]$$

subject to:

$$1\geq T_i \geq 0 \qquad \text{for } i=1,\ldots,N.$$

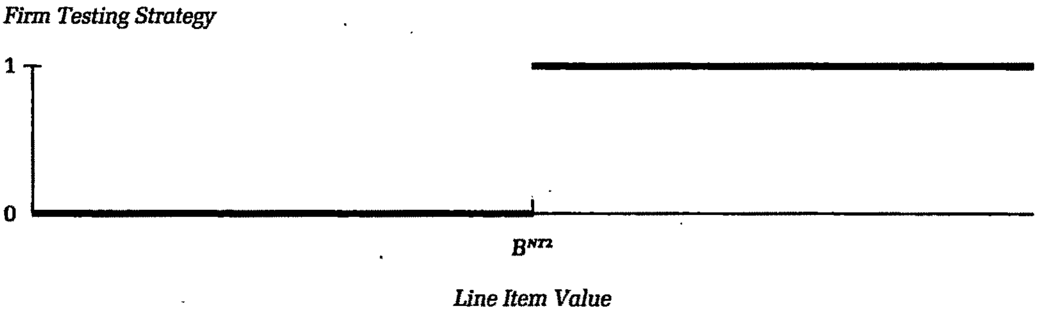
The firm's objective function is similar to that in the billing model, with the substitution of τ_i for ℓ_i.

As with the first statistical model, the solution to the firm's program is straightforward. It always tests or never tests any specific line item. The optimal rule is to test an item if and only if,

$$ap\{[1-F(0)][(hi)rh-(1-hi)\epsilon B_i]+F(0)[(lo)rl+(1-lo)\epsilon B_i]\}-C\geq 0. \tag{4}$$

Unlike the statistical billing errors model, this model connects the testing policy and item value. Because the net testing gain is larger for the more valuable items, the firm will test all items above a value cutoff, labeled B^{NT2}. This strategy is captured in figure 5.

Figure 5
The Statistical Sales Discount Model



The Strategic Sales Discount Model

The fourth model adds in the possibility of dishonest employees who may provide unearned discounts to their customers.[24] Because the discount is a set percentage of item value, the theft amount is exogenous. The size of a theft equals the size of the misappropriated sales discount ϵB_i. With this modification, the dishonest employee for line item i chooses the probability of stealing, S_i, to solve:

$$\max_{S_i} \quad S_i\{\alpha\epsilon B_i - T_i d(\epsilon B_i)[\epsilon\alpha B_i + L(\epsilon B_i)]\}$$

subject to:

$$1 \geq S_i \geq 0.$$

As before, the firm minimizes its expected losses and testing costs. It chooses a set of testing probabilities $T_i(i=1,\ldots,n)$, and solves:

$$\min_{T_1,\ldots,T_N} \quad \sum_{i=1}^{n} \Big(T_i\{\delta S_i[1-d(\epsilon B_i)]\epsilon B_i + \delta(1-S_i)a(1-p)\tau_i + (1-\delta)a(1-p)\tau_i + C\}$$
$$+(1-T_i)[\delta S_i \epsilon B_i + \delta(1-S_i)a\tau_i + (1-\delta)a\tau_i]\Big),$$

subject to:

$$1 \geq T_i \geq 0 \qquad \text{for } i=1,\ldots,N.$$

The firm's objective function is similar to that in the strategic billing errors model. Two simple changes are substituting τ_i for ℓ_i, and replacing ϵB_i for V_i. The other modification is that when an employee steals, the loss does not contain terms for statistical errors. This reflects the assumption that a statistical error and a strategic irregularity are mutually exclusive.

---

[24] Sawyer (1988, 1016–17) contains a list of 40 common frauds. One element is:

20. Recording unwarranted cash discounts.

The following theorem describes the most interesting equilibrium case.[25]

Theorem 2: If (1) the firm faces both strategic and nonstrategic errors, (2) the decline in the firm's loss when it detects a strategic irregularity is larger than the reduction when it detects a statistical error,

$$d(\epsilon B_i)\epsilon B_i+a\{[1-F(0)](-\epsilon B_i)+F(0)\epsilon B_i\}-ap_{T_i}>0,$$

(3) the firm would not test only to reduce statistical error,

$$C>ap_{T_i},$$

and (4) $B^{S|T}>B^{NT}$, then the equilibrium firm testing probabilities, employee theft probabilities, and theft amounts can be partitioned into the following regions.

1. If $B^{NT}>B_i$, then the firm never tests, a dishonest employee always steals, and he steals the amount ϵB_i.
2. If $B^{S|T}>B_i>B^{NT}$, then the firm tests the item with probability,

$$T_i=\frac{\alpha\epsilon B_i}{d(\epsilon B_i)[\alpha\epsilon B_i+L(\epsilon B_i)]},$$

a dishonest employee steals with probability,

$$S_i=\frac{C-ap_{T_i}}{\delta d(\epsilon B_i)\epsilon B_i+\delta a\{1-F(0)](-\epsilon B_i)+F(0)\epsilon B_i\}-\delta ap_{T_i}},$$

and when he steals, he steals amount ϵB_i.
3. If $B_i>B^{S|T}$, then the firm always tests the line item, a dishonest employee always steals, and he steals amount ϵB_i.

The constants $B^{S|T}$ and B^{NT} solve:

$$\alpha\epsilon B^{S|T}-d(\epsilon B^{S|T})[\alpha\epsilon B^{S|T}+L(\epsilon B^{S|T})]=0, \text{ and}$$

$$\delta d(\epsilon B^{NT})\epsilon B^{NT}+\delta a\{[1-F(0)](-\epsilon B_i)+F(0)\epsilon B_i\}=C-(1-\delta)ap_{T_i}.$$

Proof: The proof is similar to that of Theorem 1 and is available from the author.

There are several cutoffs that characterize the equilibrium behavior. The value $B^{S|T}$ is the value for which the criminal is indifferent between stealing and not stealing given that the firm always tests the item. If the value of the item is greater than $B^{S|T}$, then the dishonest employee strictly prefers to steal.

The value B^{NT} is the value for which the firm is indifferent between testing and not testing given that the dishonest employee always steals the entire item when the firm does not test. If the value lies below B^{NT} the firm never tests.

The next step is to illustrate the player's behavior in equilibrium. Recall that in this model all parameters are constant except the value of the line item. Figure 6 shows the firm's and dishonest employee's strategies.

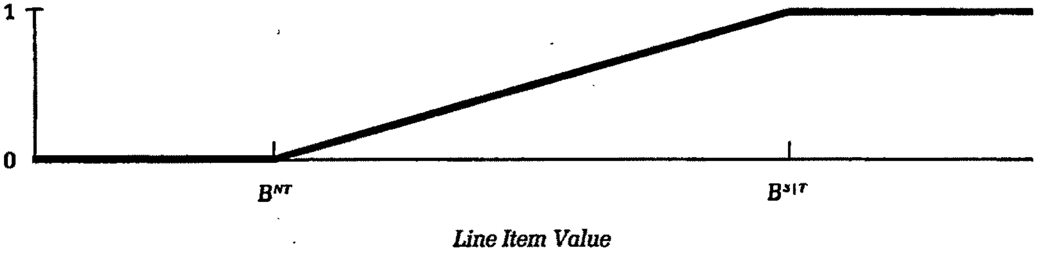
The firm's behavior is very intuitive. If all other factors are held constant, the

---

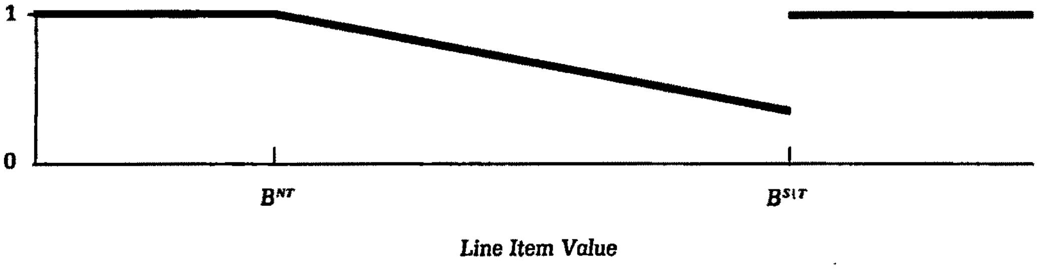
[25] The remaining possible behaviors are provided in appendix B. Both cases 2 and 3 contain only pure strategies for the firm and dishonest employees. The firm either tests or does not test an item, while the dishonest employee always steals or does not steal an item.

Figure 6
The Strategic Sales Discount Model

Panel A. Firm Testing Strategy:



Panel B. Dishonest Employee Theft Strategy:



probability that a firm tests an item (weakly) increases in the value of the line item. The firm never tests the low value line items, and always investigates the high value ones. The probability that an intermediate value item is investigated increases with the value of the line item.

The dishonest employee's strategy is similar to the firm's. The employee always steals from the highest value items because of the potential payoff. He steals from the lowest value items because he will not be caught. The probability that he steals intermediate value items decreases in the value of the item.[26]

One subtle result is that the correlation between the probability of a dishonest employee and the probability of a statistical error does affect the solution. A higher probability of statistical error lowers the value cutoffs between the equilibrium regions. All regions shift left. If the probability of a statistical error is high enough, the equilibrium shifts to a region where the firm always tests to reduce statistical error and employees always steal. In this sense, the correlation between the probability of statistical errors and dishonest employees can lead to increased theft.

---

[26] The discontinuity in the employee's theft probability reflects the switch from pure to mixed strategies. The transition point B^{SIT} is determined by the firm's profits. At B^{SIT} the firm switches from being indifferent between testing and not testing, to having a strict preference for testing. This leads to the smooth transition for the firm's strategy, but induces a jump in the dishonest employee's.

The most interesting region of the equilibrium is where the firm chooses a randomized testing strategy. In this region the firm's probability of testing increases as the value of the line item grows. This observation immediately suggests a comparison with Dollar Unit Sampling.

Corollary 2. In the strategic sales discount model, if statistical errors alone would not lead to testing, then the firm uses a testing strategy similar to Dollar Unit Cell Width Sampling (Leslie et al. 1980). It never tests low value items, always tests high value items, and randomizes over intermediate items with a probability roughly proportionate to the value of the item.

Proof: In Dollar Unit Cell Width Sampling, the firm randomly samples line items within a specific range (cell width). The firm samples one item from each cell and assigns each item *i* in the cell the expected probability,

$$T_{\text{dollar unit}} = \frac{B_i}{B_i + \text{value of all other line items in cell}} = \frac{B_i}{B_i + \text{constant}}.$$

The Dollar Unit Sampling probability can be compared to the firm's randomized testing strategy in my model, which is given by:

$$T_i = \frac{\alpha\epsilon B_i}{d(\epsilon B_i)[\alpha\epsilon B_i + L(\epsilon B_i)]}.$$

If the detection probability is very high, and the liability does not change dramatically with an increase in the amount of theft, then the firm's randomized strategy is given by:

$$T_i \sim \frac{B_i}{B_i + \text{constant}}.$$

The tie between the randomized strategies and item value is a consequence of the error and theft technology. Since both the firm and employee's payoffs are directly linked to the item value, any randomized strategies must also change with the item values.

IV. Discussion

This paper has investigated how the relationship between the firm's testing strategy and item value changes under different conditions. Two features have been varied across the four models: the link between the net gain from testing and item value, and the presence of dishonest employees. Permuting the two features yields two sets of results.

One set illustrates that the mapping between the net gain from testing and item value determines the connection between the firm's testing strategy and item value. I investigate two specific factors that affect this mapping: the accounting system itself, and the presence of dishonest employees. When the accounting transaction has no link between testing gain and item value, as in the statistical billing errors model, the item value does not have a consistent effect on the firm's strategy. The firm may test only high value items, only low value items, or any other combination of items. When there is a clear mapping between net testing gain and item value, as in the statistical sales discount model, there is a connection between item strategy and value. The firm always tests the high value items and never tests the low ones. A second method of connecting the firm's strategy and item value is to allow for the possibility of strategic, dishonest employees, as in the strategic billing errors model. In this model, dishonest employee's greed leads them to steal the entire item if they know that the firm will not investigate that item. Therefore, the firm's losses from not testing now depend on the item value, and the firm never tests low value items.

The second set of results examines the impact of introducing a strategic player. I look at how the firm responds to intelligent theft by dishonest employees. As in prior strategic auditing work, the firm may respond to strategic theft by using randomized testing strategies. The firm may use a probability of testing in order to counteract strategic theft by the dishonest employees. However, unlike previous work, my model generates randomized strategies that are related to common audit techniques. When the satiation theft amount is unrelated to item value, the firm's strategy is similar to Physical Units Sampling. When the dishonest employee's thefts are a percentage of the item value, Dollar Units Sampling emerges.

My analysis contains three major limitations. The first one is the assumption of many independent, risk neutral employees who can each perform only one theft. A more complete theory would incorporate risk aversion and allow employees to steal from multiple items. The second limitation is that I assume that the internal auditor has incentives aligned with the firm. In practice, motivating the internal auditor is a complex problem and will affect the feasible testing strategies. The third limitation is that I assume that billing errors and sales discounts can be examined separately. Typically an accounts receivable clerk can both send bills and generate discounts. If an employee performs multiple tasks, the optimal testing strategy will be more complex.

Although there is a similarity in form, the randomization in my model and in common audit sampling techniques are based on different premises. In practice, internal auditors use Dollar Unit or Physical Unit Sampling as a response to limited resources. The firm is attempting to study a population. It has a finite amount of time and money and wishes to acquire a sample that will allow it to draw inferences about the population. Dollar Unit and Physical Unit Sampling are particular ways of drawing the sample. In my model, randomization plays a different role. Here the firm randomizes even though it does not have a budget constraint. Randomization is the minimum cost method of preventing dishonest employees from stealing. In my model a firm might randomize even if there were only one item. In contrast, a necessary condition for randomization in standard audit sampling is the presence of a large number of items.[27]

Although randomization serves a different purpose in my model, there are several implications for practice. My model shows that a beneficial side effect of randomized sampling is that it may deter theft. By choosing a randomized sample, the firm makes the dishonest employee uncertain about which items will be tested, and this confusion may lead to lower levels of theft. The second implication is that the common practice of stratification into groups may still be optimal even when there may be dishonest employees. Standard practice is to group items by value, physical characteristics, and susceptibility to theft and misreporting. My model supports this practice. For example,

---

[27] Sawyer (1988, 408) states that Dollar Units Sampling should only be performed when there are more than 2,000 items.

Theorem 2 groups items into three strata, distinguished by the strategic behavior of the dishonest employees.

My paper indicates the importance of the accounting context in strategic auditing models. I show that the accounting and theft technologies have an important impact on the firm's strategy. Future research could model different sectors of the firm and generate the appropriate strategies for, say inventory. Additional "micro" models might indicate common features in the testing strategies across sectors. These commonalities could lead to a more general model of the impact of strategic play on the firm.

Appendix A

The Remaining Cases of Theorem 1 and its Proof

Case 1 has been presented in the text.

Case 2

If $\alpha V^{*1}-d(V^{*1})[\alpha V^{*1}+L(V^{*1})]<0$ and $C^{RS}>C^{T|S}$, then the equilibrium firm testing probabilities, employee theft probabilities, and theft amounts can be partitioned into the following regions.

1. If $C^{AT}>C_i$, then the firm always tests the line item, and the employee never steals.
2. If $C^{NT}>C_i>C^{AT}$, then the firm tests the item with probability,

$$T_i=\frac{\alpha}{d'(V^{*T})[\alpha V^{*T}+L(V^{*T})]+d(V^{*T})[\alpha+L'(V^{*T})]},$$

   a dishonest employee always steals, and the size of the theft, V^{*T}, solves:

$$\delta_i d(V^{*T})\Big(V^{*T}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\Big)=C_i-(1-\delta_i)a_i p_i \ell_i.$$

3. If $C_i>C^{NT}$, then the firm never tests the line item, a dishonest employee always steals, and he steals the total line item.

Case 3

If $\alpha V^{*1}-d(V^{*1})[\alpha V^{*1}+L(V^{*1})]>0$, then the equilibrium firm testing probabilities, employee theft probabilities, and theft amounts can be partitioned into the following regions.

1. If $C^{T|S}>C_i$ then the firm always tests the line item, the employee always steals and chooses theft amount V^{*1}.
2. If $C^{NT}>C_i>C^{T|S}$, then the firm tests the item with probability,

$$T_i=\frac{\alpha}{d'(V^{*T})[\alpha V^{*T}+L(V^{*T})]+d(V^{*T})[\alpha+L'(V^{*T})]},$$

   a dishonest employee always steals, and the size of his theft, V^{*T}, solves:

$$\delta_i d(V^{*T})\Big(V^{*T}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\Big)=C_i-(1-\delta_i)a_i p_i \ell_i.$$

3. If $C_i>C^{NT}$, then the firm never tests the line item, a dishonest employee always steals, and he steals the total line item.

Proof

Since the proofs for all cases are similar, I only provide the proof for the most complex case, Case 1. The proof consists of a sequence of lemmas that generate the equilibrium for the various regions. The firm and the dishonest employee do not have budget constraints, and therefore the analysis is done item by item.

The first lemma describes the firm's pure strategies.

Lemma A1: The firm has the following strategies:

1. When a dishonest employee never steals, the firm always investigates an item with cost below C^{AT}.
2. When the dishonest employee always steals the entire item, the firm never investigates an item with cost above C^{NT}.

Proof:

1. Inserting $S_i=0$ into the firm's minimization problem shows that it will always test if,

$$a_i(1-p_i)\ell_i+C_i<a_i\ell_i.$$

The cost that makes this an equality is C^{AT}.

2. Inserting $S_i=1$ and $V_i=B_i$ into the firm's minimization problem shows that it will never test if,

$$\delta_i[1-d(B_i)]\Big(B_i+a_i\{[1-F_i(0)](-O_i)]+F_i(0)U_i]\Big)+(1-\delta_i)a_i(1-p_i)\ell_i+C_i$$
$$>\delta_i\Big(B_i+a_i\{[1-F_i(0)](-O_i)]+F_i(0)U_i]\Big)+(1-\delta_i)a_i\ell_i.$$

The cost that makes this an equality is C^{NT}.

The second lemma provides the dishonest employee's pure strategies.

Lemma A2: A dishonest employee has the following pure strategies:

1. If the firm always tests the item, then the dishonest employee will never steal.
2. If the firm never tests the item, then the dishonest employee will always steal, and he will steal the maximum amount.

Proof:

1. By assumption $0>\alpha V^{*1}-d(V^{*1})[\alpha V^{*1}+L(V^{*1})]$.
2. This is the obvious solution to the dishonest employee's maximization problem with $T_i=0$.

Combining Lemmas A1 (1) and A2 (1) yields region 1. Combining Lemmas A1 (2) with A2 (2) yields region 4.

The other regions involve more complex behaviors. If $C^{NT}>C_i>C^{AT}$, then no pure strategy equilibrium exists. There are two mixed strategy equilibria which are described in the following lemmas.

Lemma A3: If $C^{RS}>C_i>C^{AT}$, then the mixed strategy equilibrium satisfies the following equations:

$$T^{RS}=\frac{\alpha}{d'(V^{RS})[\alpha V^{RS}+L(V^{RS})]+d(V^{RS})[\alpha+L'(V^{RS})]}, \tag{A1}$$

$$S_i=\frac{C_i-a_ip_i\ell_i}{\Big[\delta_i d(V^{RS})\Big(V^{RS}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\Big)-\delta_i a_i p_i \ell_i\Big]}, \text{ and} \tag{A2}$$

$$\alpha V^{RS}-T^{RS}d(V^{RS})[\alpha V^{RS}+L(V^{RS})]=0. \tag{A3}$$

Proof: Equation (A1) is a rearrangement of the employee's first order condition for the choice of the optimal amount of theft evaluated at V^{RS}.

A necessary condition for the firm to play a mixed strategy is that it is indifferent between testing and not testing; that is,

$$\delta_i S_i[1-d(V_i)]\Big(V_i+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\Big)+\delta_i(1-S_i)a_i(1-p_i)\ell_i$$
$$+(1-\delta_i)a_i(1-p_i)\ell_i+C_i=\delta_i S_i\Big(V_i+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\Big)$$
$$+\delta_i(1-S_i)a_i\ell_i+(1-\delta_i)a_i\ell_i.$$

Inserting in the value V^{RS} and solving for S_i yields equation (A2).

A necessary condition for the dishonest employee to play a mixed strategy is that he is indifferent between stealing and not stealing. His indifference condition is equation (A3).

One solution to equations (A1), (A2), and (A3) is for T^{RS}, V^{RS} to be constant, and S_i to increase with the costs of testing.

The bounds for this equilibrium are determined by equation (A2). If $C_i=C^{AT}$, then $S_i=0$. If $C_i=C^{RS}$, then $S_i=1$, where:

$$C^{RS}=\delta_i d(V^{RS})\Big(V^{RS}+a_i\{[1-F_i(0)](-O_i)+F_i(0)U_i\}\Big)+(1-\delta_i)a_i p_i \ell_i.$$

Lemma A4: If $C^{NT}>C_i>C^{RS}$, then the mixed strategy equilibrium satisfies the following equations:

$$T_t=\frac{\alpha}{d'(V^{*T})[\alpha V^{*T}+L(V^{*T})]+d(V^{*T})[\alpha+L'(V^{*T})]}, \tag{A4}$$

$$S_t=1, \text{ and} \tag{A5}$$

$$\delta_t d(V^{*T})\left(V^{*T}+a_t\{[1-F_t(0)](-O_t)+F_t(0)U_t\}\right)=C_t-(1-\delta_t)a_t p_t \ell_t. \tag{A6}$$

Proof: Equation (A4) is a rearrangement of the employee's first order condition for the choice of the optimal amount of theft.

A necessary condition for the firm to play a mixed strategy is that it is indifferent between testing and not testing. Inserting $S_t=1$ into the firm's minimization problem yields the condition,

$$\delta_t[1-d(V^{*T})]\left(V^{*T}+a_t\{[1-F_t(0)](-O_t)+F_t(0)U_t\}\right)+(1-\delta_t)a_t(1-p_t)\ell_t+C_t$$
$$=\delta_t\left(V^{*T}+\alpha_t\{[1-F_t(0)](-O_t)+F_t(0)U_t]\right)+(1-\delta_t)a_t\ell_t,$$

which can be rearranged to yield equation (A6).

In order to establish that this is an equilibrium, I need to prove that $S_t=1$. This is accomplished by showing that if T_t and V^{*T} solve equations (A4) and (A6), then the employee's profits increase in cost of testing. Since the employee's profits are zero at C^{RS}, this implies that the employee's profits are strictly positive for $C_t>C^{RS}$ and the employee will strictly prefer to steal.

Lemma A5: If $C^{NT}>C_t>C^{RS}$, then,

$$\frac{d\text{ employee's profits}}{dC}>0.$$

Proof: This proof contains four steps.

1. The second order condition for the employee's optimal amount of theft is:

$$-T\{d''(V^{*T})[\alpha V^{*T}+L(V^{*T})]+2d'(V^{*T})[\alpha+L'(V^{*T})]+d(V^{*T})L''(V^{*T})\}<0. \tag{A7}$$

2. Totally differentiating equation (A6) with respect to C gives,

$$[\delta_t d'(V^{*T})\left(V^{*T}+a_t\{[1-F_t(0)](-O_t)+F_t(0)U_t\}\right)+\delta_t d(V^{*T})]\frac{dV}{dC}=1,$$

which implies that,

$$dV/dC>0. \tag{A8}$$

3. Totally differentiating the employee's first order condition for optimal amount of theft with respect to C provides,

$$-\{d'(V^{*T})[\alpha V^{*T}+L(V^{*T})]+d(V^{*T})[\alpha+L'V^{*T})]\}\frac{dT}{dC}$$
$$-T\{d''(V^{*T})[\alpha V^{*T}+L(V^{*T})]+2d'(V^{*T})[\alpha+L'(V^{*T})]+d(V^{*T})L''(V^{*T})\}\frac{dV}{dC}=0. \tag{A9}$$

Using equations (A7), (A8), and the fact that $T_t>0$ requires $\{d'(V^{*T})[\alpha V^{*T}+L(V^{*T})]+d(V^{*T})[\alpha+L'(V^{*T})]\}>0$, implies that,

$$dT/dC<0. \tag{A10}$$

4. Totally differentiating the employee's profits with respect to C and using the envelope theorem,

$$\frac{d\text{ employee profits}}{dC}=-\frac{dT}{dC}d(V^{*T})[\alpha V^{*T}+L(V^{*T})].$$

Using equation (A10) generates the result.

The bounds for the equilibrium are easily determined. If $C=C^{NT}$, then equation (A6) does not hold. If $C=C^{RS}$, then the employee has zero profits and equation (A5) does not hold.

Appendix B

The Remaining Cases of Theorem 2

Case 2

If $C > ap_{T_i}$ and $B^{NT} > B^{S|T}$, then the equilibrium firm testing probabilities, employee theft probabilities, and theft amounts can be partitioned into the following regions.

1. If $B^{NT} > B_i$, then the firm never tests the line item, a dishonest employee always steals, and he steals the amount ϵB_i.
2. If $B_i > B^{NT}$, then the firm always tests the line item, a dishonest employee always steals, and he steals the amount ϵB_i.

Case 3

If $C < ap_{T_i}$, then the equilibrium firm testing probabilities, employee theft probabilities, and theft amounts can be partitioned into the following regions.

1. If $B^{S|T} > B_i$, then the firm always tests the item, and the employee never steals.
2. If $B_i > B^{S|T}$, then the firm always tests the line item, a dishonest employee always steals, and he steals the amount ϵB_i.

References

Anderson, U., and R. Young. 1988. Internal audit planning in an interactive environment. *Auditing: A Journal of Practice & Theory* 8 (Fall): 23–42.

Bologna, G., and R. Lindquist. 1987. *Fraud Auditing and Forensic Accounting*. 1st ed. New York, NY: John Wiley and Sons.

Dworin, L., and R. Grimlund. 1984. Dollar unit sampling for accounts receivable and inventory. *The Accounting Review* 59 (April): 218–41.

Fellingham, J., and P. Newman. 1985. Strategic considerations in auditing. *The Accounting Review* 60 (October): 634–50.

——, ——, and E. Patterson. 1989. Sampling information in strategic audit settings. *Auditing: A Journal of Practice & Theory* 8 (Spring): 1–21.

Hollinger, R., and J. Clark. 1983. *Theft by Employees*. Lexington, MA: Lexington Books.

Johnson, J., R. Leitch, and J. Neter. 1981. Characteristics of errors in accounts receivable and inventory audits. *The Accounting Review* 56 (April): 270–93.

Leslie, D., A. Teitlebaum, and R. Anderson. 1980. *Dollar Unit Sampling, A Practical Guide for Auditors*. Toronto, Canada: Pitman Publishing, Ltd.

Menzefricke, U. 1984. Using decision theory for planning audit sample size with dollar unit sampling. *The Journal of Accounting Research* 22 (Autumn): 570–87.

Newman, P., J. Park, and R. Smith. 1991. Reporting strategies in a multi-location audit environment. Working paper, The University of Texas at Austin.

Patterson, E. 1990. Stategic sample size choice in auditing. Working paper, The University of Michigan at Ann Arbor.

Sawyer, L. 1988. *Sawyer's Internal Auditing*. Altamonte Springs, FL: The Institute of Internal Auditors, Inc.

Shibano, T. 1990. Assessing audit risk from errors and irregularities. *The Journal of Accounting Research* 28 (Supplement): 110–40.

Teitlebaum, A., and C. Robinson. 1975. The real risks in audit sampling. *The Journal of Accounting Research* 13 (Supplement): 70–91.

Wallace, W. 1984. *Handbook of Internal Controls*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 346–366

Perceived Auditor Quality and the Earnings Response Coefficient

Siew Hong Teoh
University of California–Los Angeles
T. J. Wong
University of Maryland

SYNOPSIS AND INTRODUCTION: An auditor's reputation lends credibility to the earnings report that he audits. An unresolved issue is whether auditor size is correlated with auditor quality, where a high-quality auditor is defined as one who brings about more credible earnings reports. According to basic intuition and a modified Holthausen-Verrecchia (1988) model, investors' response to an earnings surprise will depend on the perceived credibility of the earnings report. In this study, we examine whether the earnings response coefficient (*ERC*) differs between Big Eight (B8) and non-Big Eight (NB8) audited firms. This provides a test of the joint hypotheses that auditor size is a proxy for auditor credibility and of the modified H-V model.

Consistent with the joint hypotheses, we find that the *ERC*s of Big Eight clients are statistically significantly higher than for non-Big Eight clients. The result obtains in both a matched sample of firms paired according to industry membership, and a switch sample of firms grouped according to shifts from and to B8 and NB8 auditors. Furthermore, the result is robust with respect to the inclusion of other explanatory factors for *ERC* that have been suggested by previous studies: growth and persistence, risk, firm size, and predisclosure information environment.

Key Words: *Auditor quality, Auditor size, Earnings quality, Earnings response coefficient, Precision of information.*

Data Availability: *Analysts' forecasts may be requested from the Institutional Brokers Estimate System (IBES); all other data are from public sources.*

We thank the anonymous referees for helpful suggestions. We are also grateful for useful comments from V. Bernard, T. Cameron, P. Hughes, J. McKeown, B. Miller, W. Teets, S. Titman, from workshop participants at UCLA, UBC, the University of Maryland, the University of Oregon, and Pennsylvania State University, and from conference participants at the 1991 AAA meeting in Nashville, TN. We thank IBES for providing analyst forecast data, and S. H. Teoh appreciates financial support from the KPMG Peat Marwick Foundation.

Submitted August 1990.
Accepted October 1992.

THE remainder of the article is organized as follows. In section I, we present the motivation for the study by explaining the hypothesized relation between investor response to earnings surprises and auditor type (B8 and NB8). In section II, we provide a simple analytical formula that relates stock price response to the precision of the earnings signal. The selection criteria for the industry-matched and the switch samples are explained in section III. The test procedures for both samples are described in section IV, which also contains a description of the control variables in the regression equation and the empirical proxies for the earnings surprise and stock price response. Results are reported in section V and section VI contains our conclusions.

I. Motivation

Recent research in the capital markets area has focused on the determinants of the *ERC*, which is the multiple that correlates unexpected earnings with abnormal changes in stock prices in response. The *ERC* is commonly estimated as the slope coefficient in a regression of the abnormal stock returns on a measure of earnings surprise. It is therefore a measure of the extent to which new earnings information is capitalized in the stock price. Current evidence indicates that the *ERC* varies cross-sectionally and intertemporally; for example, see Kormendi and Lipe (1987), Collins and Kothari (1989), Easton and Zmijewski (1989), Biddle and Seow (1991), and Lipe (1990). Collectively, these studies have shown that the *ERC* varies with the degree of persistence in earnings, the predictability of earnings, the security's covariation with the market return, growth opportunities of the firm, and industry membership.

Researchers in this area have begun to examine whether the stock price reaction to earnings surprises is related to the quality of the reported earnings numbers. Imhoff and Lobo (1992) find that firms with low consensus in the analysts' forecasts of earnings tend to have a low *ERC*. Although it is possible that high prior uncertainty about the underlying value of the firm would also increase the dispersion in forecasts, Imhoff and Lobo conclude that it is more likely to proxy for the noise in accounting measures than for any prior uncertainty about underlying cash flows.

Lang and McNichols (1990) use the correlation between cash flow and earnings as a proxy for earnings report quality. To the extent that high-quality auditors assure a greater conformance of the financial report with generally accepted accounting principles (GAAP), less discretion in the management of accruals is permitted. Therefore, one might expect a greater correlation between expected future cash flows and accounting earnings with high-quality auditors.[1]

In a similar spirit, we expect earnings reports to have a greater effect on investors' valuation of the firm when the reported numbers more accurately reflect true economic value. This can be demonstrated theoretically with a simplified version of the Holthausen and Verrecchia (1988) model presented in the next section. The model predicts that the magnitude of the stock price response increases with the precision (inverse of the variance of the noise term) of the information.

---

[1] This is not, however, a logically necessary consequence. To give an extreme example, an auditor who incorrectly required the firm to record zero accruals always would cause a perfect correlation between cash flows and earnings. More generally, for a good auditor, the divergence of earnings from cash flows may capture meaningful economic events, in which case a lower correlation of earnings and cash flows could be associated with more informative earnings. Lang and McNichols (1990) find that the ERC is lower for firms with higher correlation between net income and contemporaneous cash from operations, and for firms with higher accruals, contrary to their predictions.

Since investors cannot directly observe the underlying true earnings of the firm, they have to rely on reported accounting numbers. To safeguard the credibility of these reported figures, external auditors must certify that they conform to GAAP, which assures investors of the reliability of financial data. This reflects the attestation role of auditing (see, e.g., Abdel-khalik and Solomon 1988). A more skillful auditor will presumably be able to bring closer concordance of the reported earnings with GAAP.[2] It seems plausible, therefore, that if the auditor's quality (skill) is perceived by investors as high, they will respond strongly to surprises in reported earnings.

Although a link between informativeness of earnings and the firm's conformance with GAAP is plausible, it is not crucial for the argument. So long as some auditors are perceived to follow policies that cause reported earnings to be more informative about value than other auditors, valuation theory predicts that the *ERC* will be different for different auditors. An auditor's "quality" can then be defined as the characteristic leading to greater informativeness of reported earnings.[3] Whether some auditors do in fact produce earnings reports with greater credibility for investors is an interesting issue. To examine this, we focus on an observable characteristic of the audit firm size, which managers have argued is associated with audit quality.

DeAngelo (1981) suggests that the initial client-specific setup costs confer economic rents to the incumbent auditor in later periods as a consequence of a first-mover advantage. Because a failure to detect any material omissions and misrepresentations in one client's records may cause the auditor to lose these rents from some or all clients, a larger audit firm has a greater commitment to accuracy. Dopuch and Simunic (1980, 1982) propose that investors may rationally perceive B8 auditors as of higher quality because larger auditors have more of the observable characteristics associated with quality (i.e., specialized training, accreditation by some reputable agency, and peer reviews). In a model on the determinants of optimal size and structure of the auditing firms, John (1991) shows that auditor quality increases with size.

A survey of management by Carpenter and Strawser (1971) concludes that the majority of firms issuing equity stock switch to B8 auditors to secure the best price for the shares. Other empirical studies that support the correlation of size and quality include Palmrose (1988) and Beatty (1989). Using the incidence of litigation as a measure of auditor quality, Palmrose finds that the NB8 auditors as a group experience more litigation than B8 auditors. This result holds despite the possibility of a higher tendency to sue larger auditors who have deeper pockets. Beatty finds that the initial return on an initial public offering is higher for NB8 than for B8 client firms (i.e., there is less underpricing for B8 client firms).

---

[2] This does not imply that the financial reports audited by low-quality auditors necessarily violate GAAP, but that the likelihood and amount of deviation of the reported numbers from true economic levels are lower for high-quality auditors.

[3] The literature on audit quality recognizes that the audit service has multiple attributes; for example, Simunic and Stein (1987) characterize the service in terms of (1) control (provision of reliable information for internal control purposes), (2) product line (availability of consulting services as well as the traditional audit service), and (3) credibility. In a study of private companies that voluntarily hire external auditors, O'Keefe and Barefield (1985) found that the majority (57 percent) cited enhanced credibility with outside parties as the major reason for the hire. We focus on credibility because of its importance, and because the other characteristics would seem to have relevance for the *ERC* primarily via their effects on the credibility of the financial statements.

Conversely, some studies have not found an association between size of the audit firm and measures of quality. Imhoff (1988) surveyed financial analysts and found that they perceive no differences in quality between B8 and NB8 clients. Event studies of auditor switches provide evidence that is hard to assess. Nichols and Smith (1983) find that the stock price reaction to an announcement of a change from a NB8 to a B8 auditor is larger than for a change from B8 to a NB8 auditor, but the difference is not statistically significant. They hypothesize that a higher-quality auditor will be more credible, and hence a switch to a B8 auditor should be good news. A theoretical justification for the link between firm value and auditor quality is provided by Titman and Trueman (1986) and Datar et al. (1991). However, these models analyze first-time selection of auditors for a firm that is going public, and therefore their prediction relating auditor quality to value does not necessarily apply to auditor switches. For example, Teoh (1992a) shows that auditor switches can convey either favorable or adverse information about firm value to investors even if all auditors are identical. Hence, a more complex model that considers the issues associated with switching per se as well as the quality of the chosen auditor is needed before one can predict whether cross-quality switches are good or bad news.[4]

This research differs from previous studies in that it examines the link between auditor size and a crucial ultimate output of the auditor, the credibility of the financial report. Important consequences related to auditor size may pertain to litigation and information revelations, but we believe it is also important to examine whether large auditors are more credible on the level of financial reporting. The empirical test considers the relation betweeen auditor size and the *ERC*. It is therefore a joint test of two hypotheses: that earnings report quality is an important determinant of the *ERC* and that auditor credibility increases with auditor size. If one maintains that auditor size and quality of the financial report are related, then the evidence provided here concerns whether the quality of reported earnings is an important determinant of the *ERC*. If one maintains that the *ERC* increases with earnings quality, then the evidence indicates whether auditor credibility increases with auditor size.

II. Precision of the Earnings Signal and the *ERC*

In this section, we develop a simple analytical formula that relates the stock price response to the precision of the earnings signal. We adapt a single-period with a single information signal model from Holthausen and Verrecchia (1988). Our aim is to provide the simplest setting sufficient for a test of the relation between the market response to an information signal and the precision of the signal. Some strong simplifying assumptions will be made, but we expect the basic results to hold in more general settings with multiple information signals. Therefore, we abstract from some pertinent factors that would be incorporated in a more general model, such as the firm's riskiness, the degree of earnings persistence and predictability, and other time series characteristics.[5]

---

[4] Using a larger sample of switches, Teoh (1992b) finds that the stock price reaction to cross-quality switches depends on the pre-switch audit opinion.

[5] For a more general analysis that incorporates these issues in deriving the theoretical measure of the *ERC*, see Kormendi and Lipe (1987) and Lipe (1990).

At the start of the period at date 0, the price of a firm *i* is P^0. (Since all parameters in the model refer to firm *i*, we do not use the subscript in this section for simplicity.) At the earnings announcement date 1, the price of the firm is P^1. The stock price response to the announcement is measured by $\delta = P^1 - P^0$. Let the unknown random value of firm *i* be $\tilde{u}$ and assume it is normally distributed with mean *m* and variance *v*.[6] At date 0, the price of firm *i* is:

$$P^0 = E[\tilde{u}] = m. \tag{1}$$

Let x be the earnings signal that is released at date 1, and let the earnings signal communicate the true value of the firm with noise, $\tilde{\varepsilon}$: $\tilde{x} = \tilde{u} + \tilde{\varepsilon}$. The random variable $\tilde{\varepsilon}$ has normal distribution with mean zero and variance η. Hence, $x - m$ is a measure of the earnings surprise for firm i. At date 1:

$$P^1 = E[\tilde{u} \mid \tilde{x} = x] = m + \frac{v}{v+\eta}(x-m). \tag{2}$$

Therefore, the stock price response is:

$$\delta = P^1 - P^0 = \frac{v}{v+\eta}(x-m), \tag{3}$$

which is a positive linear function of the measure of earnings surprise. The *ERC* is the ratio $v/(v+\eta)$. It is clear that the higher the earnings signal x, the more favorable the investor response; that is, $\partial\delta/\partial x = (v/v+\eta) > 0$. More interestingly, the *ERC* will vary with the amount of prior uncertainty, *v*, and the noise in the earnings signal η. This can be seen by differentiating the *ERC* with respect to *v* and η, respectively:

$$\frac{\partial ERC}{\partial v} = \frac{\eta}{(v+\eta)^2} > 0, \tag{4}$$

$$\frac{\partial ERC}{\partial \eta} = -\frac{\eta}{(v+\eta)^2} < 0. \tag{5}$$

Thus, the stock price response increases with *v* and decreases with η. Note that *v* and η measure different types of uncertainty; *v* measures the prior uncertainty of investors about the underlying value u of the firm, whereas η measures the noise in the earnings signal x.

Equation (4) implies that the stock price response is larger for firms with high *v*, for given η, because the informational value of the earnings signal is greater when there is greater prior uncertainty.[7] In contrast, η measures the noise in the earnings signal, so that high η, for given *v*, implies a less accurate (credible) or lower quality earnings signal. Therefore, the *ERC* will increase with the quality of the earnings signal when holding constant differences in prior uncertainty about underlying cash flows, as implied by equation (5). We test this prediction by using auditor size as a proxy for earnings quality.

---

[6] The parameter *m* can also be viewed as representing the net present value of all future cash flows discounted at the appropriate rate of interest in a multiperiod setting. It would, therefore, incorporate any time series characteristics of the underlying cash flow process.

[7] Note that high *v* does not imply low predictability in the sense used by Lipe (1990, fn. 9).

III. Sample Selection

In the empirical test, we estimate a cross-sectional multiple regression model of abnormal stock returns on a measure of earnings surprise using ordinary least squares (OLS), and test for a difference in the *ERC* between the B8 and NB8 groups. Other explanatory variables included in the regression equation control for other determinants of the *ERC* that have been identified in previous studies.

The regression is estimated with two alternative samples of firms obtained from the COMPUSTAT industrial tapes, which include firms listed on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and NASDAQ (National Association of Security Dealers Automated Quotations). The first sample contains industry matched pairs of firms to control for differences in the information environment. A NB8 client firm is matched with the B8 client firm closest in size in the same industry. The second sample contains firms that have switched across auditor class groups, so the firm's past history acts as its control in the switch sample. The hypothesis test evaluates the difference in the *ERC* when the firm is audited by a B8 versus NB8 auditor.

We present results from both samples as a robustness check because the alternative designs control for different sets of conditions that may be relevant for the *ERC*. The matched-pair sample has the advantage of holding constant the calendar period and controlling for conditions common among firms in the same industry that might affect the *ERC*. The switch sample uses the same firm as the control, and so controls for firm-specific measurement errors in unexpected earnings and residual stock returns, as long as the distribution of these errors are the same before and after the switch.

The Matched-Pair Sample

Noise in accounting earnings, measurement error in unexpected earnings and variations in the information environment can result in differences in the estimated *ERCs* across firms, apart from the hypothesized factor of difference in auditor credibility. As shown in section II, tests need to control for differences in v, the variance of the underlying true cash flows of the firm. Since firms in the same industry face common demand and supply uncertainties, it is plausible that v will vary more among firms across different industries than within the same industry. Furthermore, if industry members engage in similar transactions and account for them with similar accounting methods, then the noise in accounting measures of true economic earnings could be related to industry membership. These factors suggest the need to match firms by industry membership. Empirical support is provided by Biddle and Seow (1991) and Teets (1992), who showed that *ERCs* vary across industries. We attempt to control for other relevant factors not captured by industry membership by including them as other explanatory variables in the regression equation, which we will describe later.

The sample is selected from the primary, secondary, tertiary (PST), and full coverage (FCI) files of the 1990 COMPUSTAT annual industrial tape, and is limited to firms with earnings information in any year in the period 1980-1989. Firms are also selected from the research file if earnings data are available during the 1981–1988 period for up to three years before deletion. We exclude the three years preceding deletion to prevent contamination of our results by circumstances related to the deletion.[8]

---

[8] For example, Lang and McNichols (1990) showed that *ERCs* might be affected by financial distress.

The PST and research files contain firms traded on major exchanges, whereas the FCI file contains NASDAQ firms. There were 12,940 firm-year observations from the PST file, 10,465 observations from the FCI file, and 4,579 observations from the research file.[9]

We use the analysts' consensus forecast as a measure of the market's expectation for earnings. Thus, we require that the firm be followed by the Institutional Brokers Estimate System (IBES) and that the most recent consensus forecast be made within 200 trading days of the earnings announcement date. These requirements result in 9,624 PST observations, 5,215 FCI observations, and 2,604 research file observations.

Next, we require that the firms have daily stock returns data available on tape from the Center for Research in Security Prices (CRSP) during the period between the IBES consensus forecast date and the earnings announcement date. This further reduces the sample to 9,264 PST observations, 4,998 FCI observations, and 2,513 research observations. Of this total sample, 15,480 observations are audited by B8 firms and 1,297 are audited by NB8 firms. The greater frequency of B8 clients is not surprising because these auditors tend to audit larger firms, and larger firms are more likely to be followed by analysts and by COMPUSTAT.

Industry matching is accomplished by first eliminating any four-digit SIC code industry if it is not represented by both B8 and NB8 client firms in our sample. Next, each NB8 firm is matched with a B8 firm in the same industry in the same fiscal year. If no B8 firm with the same four-digit SIC code can be found, the three- and then two-digit codes are used to find a match. If there are several B8 candidates, the firm with the closest total asset measure is selected.[10] Finally, we delete observations for which the absolute value of the measure of earnings surprise exceeds 100 percent (see sec. IV for a description of the earnings surprise measure). The final sample consists of 1,263 four-digit firm-year pairs of observations, four three-digit pairs, and 15 two-digit pairs.[11]

A wide spectrum of industries is covered in the matched-pair sample, and a total of 194 four-digit codes are represented. The extractive industries are the most heavily represented (e.g., gold and silver mining is 3.52 percent of the sample; crude oil and natural gas is 3.05 percent). Other major categories are the pharmaceutical (3.36 percent), telephone and telegraph (2.89 percent), and the eating establishment (2.58 percent) industries. Atiase (1985) suggested that the predisclosure information environment may be different for firms listed on different exchanges, but, as shown below, the distribution of firms across exchanges is not different between the auditor class groups. Consequently, we do not include an exchange dummy in the regression equations.

Auditor Class	*NYSE and AMEX firms*	*NASDAQ firms*
B8	778 (60.7 percent)	504 (39.3 percent)
NB8	800 (62.4 percent)	482 (37.6 percent)

[9] We rely on annual rather than quarterly data for our study because only the annual accounting numbers are required to be audited by external auditors for publicly traded firms.

[10] Since this does not fully ensure adequate size matching, size is further included as a control variable in the cross-sectional regression. See section IV for furthur discussion of the effect of size on the ERC.

[11] We also considered an industry-matched sample without the pairwise matching of firm-year observations. This sample has 11,208 observations consisting of 9,928 Big Eight firms and 1,282 non-Big Eight firms. This procedure utilizes more observations, but a disadvantage is that some industries carry relatively more weight between the auditor class groups in the cross-sectional regression equation. The results are qualitatively similar and are not reported.

Table 1
Types of Auditor Switches by Year

	Switches	
Switch Year	*Big Eight to Non-Big Eight (B-N)*	*Non-Big Eight to Big Eight (N-B)*
1973	1	3
1974	0	5
1975	1	1
1976	0	1
1977	1	0
1978	1	4
1979	0	3
1980	1	3
1981	3	3
1982	2	5
1983	0	6
1984	1	6
1985	4	8
1986	4	23
1987	1	35
1988	0	8
Total	20	114

Note: Three B-N switches and four N-B switches followed qualified opinions by former auditors.

The Switch Sample

An initial sample of 160 switches taken from Teoh (1992b) contains 59 switches from B8 to NB8 auditors (B-N) and 101 reverse switches (N-B). The switches from 1977–1985 are from *Who Audits America*, and those from 1973–1976 are from Disclosure, Inc. We supplement this sample with switches available from COMPUSTAT for 1985–1988 using the following selection criteria: (1) accounting and stock return data must be available on COMPUSTAT and CRSP tapes for the two years before and two years after the switch, (2) the switching firm must not be a subsidiary of a company, and (3) the firm must have retained the same auditor for two consecutive years preceding and following the switch.[12]

We further require that the firms have analyst forecasts of earnings during the test period. We use the IBES analyst consensus forecasts when these are available (1976 and after) and supplemented forecasts from *Value Line* for 15 switches occurring prior to 1976 (three are B-N and 12 are N-B switches). The final sample contains 134 switches (114 N-B and 20 B-N switches), and their distribution by type over the period 1973–1988 is presented in table 1.

The small number of B-N switches reflects the finding of the Public Accounting Report (February 1984) that B8 auditors have gained clients who trade on major exchanges relative to NB8 auditors. There is a trend towards more switches in more recent years, and there were considerably more N-B switches in 1986 and 1987 than in

---

[12] Note that the later switches include NASDAQ firms, whereas the initial sample contains only NYSE or AMEX firms. No firm switched auditors and moved to another exchange at the same time.

Table 2
Summary Characteristics of Matched-Pair and Switch Samples

	Mean		Median			
Variable	*Big Eight*	*Non-Big Eight*	*Big Eight*	*Non-Big Eight*	*T-statistic*	*p-value*
Matched-Pair sample:						
CAR1	−0.0229	−0.0182	−0.0087	−0.0086	−0.7761	0.4378
CAR2	−0.0212	−0.0186	−0.0081	−0.0103	−0.3977	0.6909
UE	−0.0248	−0.0218	−0.0012	−0.0018	−0.6946	0.4874
MB	2.5742	2.5377	1.5507	1.8391	0.0735	0.9414
LMV	4.4909	4.5184	4.3425	4.3811	−0.4700	0.6384
β	1.2042	1.2277	1.1652	1.1897	−1.0205	0.3076
N	5.6176	4.8293	3.0000	3.0000	3.3044	0.0010
CAR period	28.8029	29.2142	18.0000	16.0000	−0.3103	0.7564
Switch sample:						
CAR1	0.0123	−0.0247	0.0051	−0.0130	3.0627	0.0023
CAR2	0.0080	−0.0216	−0.0032	−0.0144	2.1203	0.0345
UE	−0.0201	−0.0242	−0.0012	−0.0018	0.4692	0.6391
MB	2.1604	2.0867	1.3871	1.5834	0.3299	0.7416
LMV	4.4616	4.5006	4.3530	4.3038	−0.2835	0.7769
β	1.2499	1.1923	1.1574	1.1481	1.1667	0.2439
N	6.8969	6.1734	5.0000	4.0000	1.2191	0.2233
CAR period	27.6527	30.9815	17.5000	17.0000	−1.1781	0.2393

Note: *CAR1*=cumulative market-adjusted return; *CAR2*=cumulative market model abnormal return; *UE*=unexpected earnings; *MB*=market value to book value; β=market model slope coefficient; and N=number of analysts' forecasts in the consensus forecast. The T-statistic tests the significance of the difference in the mean levels of the variables between Big Eight and non-Big Eight firms. The p-value measures the two-tailed significance level of the T-statistic.

any of the other years, which reflects the inclusion of NASDAQ switches in the later years (35 of 83 switches in 1984–1988 are by NASDAQ firms).[13]

The switch sample also encompasses a wide spectrum of industries. A total of 15 two-digit SIC codes are represented by 20 B-N switches. With at most two B-N switches in any one industry, there appears to be no clustering of the B-N switches. The 114 N-B switches comprise 44 two-digit SIC codes, and there are only three industries with more than six switches (the chemical industry, ten switches; industrial machinery and computer equipment, eight switches; and the electronic and electric industry, 13 switches). Furthermore, there is an overlap of B-N and N-B switches in 11 industries. Thus, as in the matched-pair sample, industry effects are unlikely to account for the results.

IV. Test Procedures and Data

A multiple regression of abnormal stock returns on earnings surprise and other control variables with a dummy variable for the auditor class is used. The regression procedure constrains the regression coefficients to be identical for firms within the same auditor class, but they are allowed to vary between auditor classes. For the industry-matched sample, the data is pooled across firms and years so that the following cross-sectional time-series regression is estimated with the OLS method:

---

[13] Time clustering of switches does not appear to be important for the results of the study. Similar regression results are obtained when switches in 1986 and 1987 are dropped from the sample.

$$\begin{aligned} CAR_{it} = {} & \lambda_0 + \lambda_1 D_{it} + \lambda_2 UE_{it} + \lambda_3 UE_{it} D_{it} + \lambda_4 UE_{it} MB_{it} + \lambda_5 UE_{it} MB_{it} D_{it} \\ & + \lambda_6 UE_{it} \beta_{it} + \lambda_7 UE_{it} \beta_{it} D_{it} + \lambda_8 UE_{it} LMV_{it} + \lambda_9 UE_{it} LMV_{it} D_{it} \\ & + \lambda_{10} UE_{it} \frac{1}{N_{it}} + \lambda_{11} UE_{it} \frac{1}{N_{it}} D_{it} + \epsilon_{it}, \end{aligned} \tag{6}$$

where:

CAR_{it} = cumulative abnormal return for firm i, continuously compounded between the forecast date and the earnings date;
UE_{it} = earnings surprise for firm i, as calculated in equation (7) and described later;
D_{it} = dummy variable taking a value of 1 for a non-Big Eight client, 0 otherwise;
MB_{it} = market value to book value as a proxy for growth and persistence;
β_{it} = market model slope coefficient as a proxy for firm risk;
LMV_{it} = natural log of market value as a proxy for firm size;
N_{it} = number of analysts' forecasts included in the consensus forecast as a proxy for noise in the predisclosure environment; and
ϵ_{it} = error term assumed to be distributed $N(0,\sigma_i^2)$.

The average intercept of B8 firms is measured by λ_0; λ_1 measures the shift in the intercept for NB8 firms from that for B8 firms; λ_4, λ_6, λ_8, and λ_{10} are coefficients for the *ERC* contributed by the control variables for the B8 group. The shift in the control variable coefficients for the NB8 group from the slope coefficients of the B8 group is estimated by λ_5, λ_7, λ_9, and λ_{11}. The residual *ERC* left unexplained by these control variables is estimated by λ_2 for the B8 group, and the shift of this coefficient for the NB8 group from that of the B8 group is estimated by λ_3. Thus, the residual *ERC* for the NB8 group is estimated by the sum of $\lambda_2+\lambda_3$. The hypothesis examines whether the residual *ERC* is smaller for NB8 than for B8 clients, which implies a test of whether the coefficient λ_3 is negative:

H_0: $\lambda_3 \geq 0$.
H_a: $\lambda_3 < 0$.

For the switch sample, we estimate the pooled cross-sectional regression with control variables as in equation (6), but with one modification for the dummy variable: $D_{it}=1$ if the firm is a NB8 client in year t, 0 otherwise. In this case, λ_2 measures the average residual *ERC* in the years when the firm is audited by a B8 auditor. So, for B-N switches, λ_2 measures the residual *ERC* in the years prior to the switch; for N-B switches, the years after the switch. The difference in residual *ERC* between the years when a B8 auditor is used versus the *ERC* in the years when a NB8 auditor is used is measured by λ_3. The test of whether the residual *ERC* is smaller with a NB8 auditor than with a B8 auditor is operationalized as a test for whether $\lambda_3 < 0$, as in the industry-matched sample.

The control variables for growth and persistence, firm risk, firm size, and the noise in the predisclosure environment are suggested by previous studies of the determinants of the *ERC*. These studies differ in many dimensions, such as the *ERC* determinants examined, the empirical measures for the determinants, and the test methods and cumulation periods. Kormendi and Lipe (1987) and Lipe (1990) suggest that earnings persistence and earnings predictability are important determinants. Collins and Kothari (1989) emphasize earnings growth and firm risk in addition to earnings persistence. Easton and Zmijewski (1989) consider the revision in analysts' forecasts (which they interpret as a measure of persistence) and expected rate of return (related to firm risk). Some studies have also considered the size of the firm but the evidence has been conflicting. For example, Easton and Zmijewski find firm size to be generally unimportant in determining the *ERC*, and Lipe finds it marginally significant (when measured by market value of equity). However, Shevlin and Shores (1990) report that firm size might be correlated with the other control variables discussed here, and Atiase (1985) suggests that predisclosure environments differ according to size.

B8 and NB8 clients may differ in these various dimensions because of differences in the underlying economic characteristics of the firms. If these factors are important for determining the *ERC*, a difference in the *ERC* may result even if auditor class is not relevant. Because we need to adjust for these confounding factors in the regression in order to measure the incremental effect of auditor class on the *ERC*, these other determinants are included in the regression if empirical proxies can be reasonably measured without unduly sacrificing observations (see the latter discussion on the empirical proxies). Therefore, λ_3 measures the effect of auditor class on the residual ERC that is not explained by these factors. Thus, the test procedure here will provide a conservative measure of the importance of credibility if differences in the underlying economic characteristics of the firm are related to the *ERC* by virtue of their correlations with auditor size (and hence, credibility).

Empirical Proxies

Unexpected earnings are measured as the actual earnings disclosed minus a measure of investors' prior expectation of earnings scaled by the stock price,

$$UE = \frac{Actual\ Earnings - Expected\ Earnings}{Price}, \tag{7}$$

where *Price* is the stock price on the day prior to the start of the cumulation period for stock returns, if available; otherwise, it is the fiscal year-end stock price.[14] Christie (1987) and Kormendi and Lipe (1987) have suggested that price is the appropriate scaling factor from the theoretical derivation of the ERC based on the dividend and earnings capitalization formulae.

In selecting a proxy for expected earnings, we attempt to minimize measurement error by using the IBES analyst consensus (median) earnings forecasts issued before the earnings announcement and selected for a minimal time lag between the forecast and the earnings announcement. Timely forecasts are more likely to capture the market's expectation of earnings more accurately. We discard any observation when $|UE|$ exceeds 100 percent to avoid using any observation that may have an undue influence on the regression parameter estimates (see, e.g., Collins and Kothari 1989). We also eliminated forecasts made more than 200 trading days prior to the earnings announcement date because news announcements subsequent to the forecast are likely to have resulted in the market's revising its expectation.[15]

---

[14] This prevents losing about 7 percent of observations in the matched-pair sample and 3.6 percent in the switch sample. The prices could be missing because of infrequent trading of the NASDAQ firms.

Prior evidence suggests that analyst forecasts are better proxies for expected earnings than time-series models of earnings and that more recent forecasts outperform older ones (see, e.g., Brown et al. 1985; O'Brien 1988). Analysts' forecasts are likely superior to an annual random-walk model for expected earnings for our study because our test period is shorter than a year. Unlike analysts' forecasts, a random-walk model for expected earnings would not incorporate new sources of information after the announcement of the previous year's earnings.[16]

The analyst consensus forecast of annual earnings is updated every third Thursday of the month. The most recent consensus forecast available prior to the earnings announcement is selected as follows: If the earnings announcement for a firm occurs after the third Thursday of the month, then the consensus forecast for that month is selected. Otherwise, the consensus forecast for the previous month is selected. This ensures a minimum time lag between the release of annual earnings and the most recent consensus forecast available prior to the firm's disclosure of earnings. The actual earnings numbers and dates for each firm are the earnings per share figures published in *The Wall Street Journal Index.*

The cumulative, continuously compounded abnormal return, CAR_{it}, is the empirical proxy for δ_i in equation (3). It measures the residual price change for firm i that is unexplained by the beta risk of the firm. This is presumed to be the investors' reaction to new information in the earnings announcement. We cumulate the abnormal return from the day of the consensus forecast to the day of the earnings announcement to minimize measurement error in the earnings surprise. Thus, we measure the abnormal return over the same period as the forecast horizon.[17] We estimate CAR_{it} in two ways. The first measure assumes that the market model β is unity for all firms in the sample, so the abnormal return is estimated as the raw return minus the market return of the same day, continuously compounded between the IBES consensus forecast date and the earnings announcement date:

$$CAR1_{it} = \sum_{t=IB}^{ED} \log(1 + R_{it} - R_{mt}),$$

where *IB* is the *IBES* consensus forecast date and *ED* is the earnings announcement date; the market return is the CRSP equally-weighted stock index if the firm was traded on the NYSE or AMEX, and is the NASDAQ market index for NASDAQ firms. The second measure is the residual from a market model regression and is calculated as:

---

[15] The results are similar when using a 100-day cutoff.

[16] O'Brien (1988) does note that analysts' forecasts are not necessarily best in association studies. Although a more sophisticated time-series model might be superior to the annual random-walk model, the data requirements for estimating such a model would be too restrictive for our study.

[17] A disadvantage is that the measure of abnormal return (the dependent variable) includes noise in the form of other announcements during this interval. The industry-matching design and switch-sample design, a 200-day maximum period, and large sample sizes are attempts to limit the noise introduced by other news. Alternatively, an analysis over, e.g., days −1 and 0 relative to the earnings announcement date, could be used. Unlike problem of noise in the *dependent* variable over a long period, the analysis for a short period has the serious disadvantage that the regression coefficients are biased because of measurement error in the *independent* variables. The size of the measurement error bias and the amount of correction for the bias when using the Easton and Zmijewski (1989) technique will differ between Big Eight and non-Big Eight subsamples. Since the empirical tests center on comparing regression coefficients between the groups, an analysis over a short period will in general lead to invalid inferences.

$$CAR2_{it} = \sum_{t=IB}^{ED} \log(1 + R_{it} - \hat{\alpha}_i - \hat{\beta}_i R_{mt}),$$

where the market model parameters $\hat{\alpha}_i$ and $\hat{\beta}_i$ are estimated from a time-series firm-specific regression of the stock return on the market return defined as for $CAR1$. The estimation period for this regression covers the 360 days prior to the IBES consensus forecast date. There must be a minimum of 100 daily returns in the period for the market model regression.

We consider both measures of CAR_{it} to gauge whether the results are sensitive to errors in measuring the relevant variables, since different measures will introduce different types of noise in measurement. For example, $CAR1$ will be mismeasured if the true β of the firm is not unity, and $CAR2$ may be a noisy measure if estimates of the parameters are inaccurate because of infrequent trading among NASDAQ firms in the sample (see Marais et al. 1989). Thus, the simpler method of the first measure may be preferable to avoid introducing unnecessary noise into the regression equation.

We measure MB_{it} as the market to the book value of equity at the beginning of each year t. Because it is plausible that future earnings are affected by growth opportunities, the higher the market to book value, the larger is the expected earnings growth. Collins and Kothari (1989) suggest that this ratio is also affected by persistence. To the extent that this is the case, it provides a normalization for persistence as well.[18]

We obtain β_{it} from the market model estimation for expected returns discussed earlier, and it is a proxy for the riskiness of the firm's earnings. In Kormendi and Lipe (1987), and Easton and Zmijewski (1989), among others, the firm's riskiness enters the valuation models as the present-value discount factor for the revision in the expected earnings in each period. Thus, a negative relation between the *ERC* and firm risk is predicted. The empirical evidence, however, has been mixed in these previous studies. Easton and Zmijewski did not find a significant partial correlation between the *ERC* and β. Collins and Kothari report a significant positive relation between β and their return-response coefficient in a reverse regression of the earnings surprise on a variety of explanatory variables, including the raw return, and β. They interpret this as confirming a negative relation between the *ERC* and risk.[19] Another confounding factor in the relation between the *ERC* and risk might be gleaned from the model in section II. We might think of the firm's risk measure as a proxy for v in the model; more risky firms have higher variance of the underlying cash flows. If this is true, then equation (4) predicts that the *ERC* should be positively related to risk. A more general equilibrium model is required before the relation between risk and the *ERC* can be predicted. Thus, we include β in our regression as a control variable without imposing any priors about the sign of the coefficient.

---

[18] A separate persistence measure is not included because the earnings history of many firms is insufficient for adequate empirical measure for a large number of the firms in our sample. Kormendi and Lipe (1987) used 35 consecutive earnings observations per firm, and Easton and Zmijewski (1989) used 20 in their time-series measure of persistence. Since our study emphasizes audited earnings, the requirement of 20 consecutive annual earnings prior to the test period would eliminate a considerable number of our firms, especially the NASDAQ firms.

[19] Leamer (1978) suggests caution in interpreting results of multiple regressions when reverse regression methods are employed and there is correlation among the independent variables.

We also include a measure of firm size (the natural log of market value at the beginning of the year) in the regression equation, as a control variable for other possible missing factors. (As noted earlier, the economic rationale for including size is unclear, and previous studies have not found size to be important.) Finally, we include the number of analysts' forecasts used in computing the consensus forecast in an attempt to control for differences in the predisclosure environment. If there are more analysts following the firm, there may be more information about the firm, and so the market's expectation of the firm's earnings might be more accurately estimated.

We use a binary measure for these control variables because of the possibility of nonlinearity. Thus, MB_{it}, β_{it}, LMV_{it}, and $1/N_{it}$ take on the value 1 when the measures are above the median and 0 otherwise. The binary measures are intended to reduce noise from mismeasurement of these variables, and so reduce the possibility that large measures have an undue influence on the regression.[20] Care, therefore, must be exercised in interpreting the results for these other variables.

To check whether the B8 and NB8 groups differ in the dimensions of these control variables, we provide summary statistics in table 2 for each group separately, for both the matched-pair and the switch samples.[21] The data indicate that in the matched-pair sample, the groups do not differ significantly in the mean and median levels of the proxies for growth and persistence, firm risk, and firm size. The mean number of analysts following the firms is significantly different between the groups, with B8 firms having more visibility with analysts. This result should be interpreted with caution because the median number of analysts is the same for both groups. This suggests the presence of nonnormality in the distribution of analysts, and therefore, use of a binary measure may be appropriate. The mean levels of the control variables do not differ significantly between the auditor class groups in the switch sample.

Summary statistics for the dependent variable, the earnings surprise measure, and the length of the cumulation period for the abnormal returns are also included in table 2. None of the variables differs significantly in the mean levels between the B8 and the NB8 firms in the matched-pair sample. In the switch sample, the earnings surprise and the cumulation window also do not differ between the groups, but the cumulative abnormal returns are larger when B8 auditors are used.[22] Different degrees of noise in the measure of *CAR* between the auditor groups because of different *CAR* periods could have resulted in heteroscedasticity in the errors of the regression equations. The insignificant difference in the mean and median *CAR* periods between the groups in table 2 eliminates this potential source for heteroscedasticity in the regression equations. In addition, the Pearson correlation matrix for the variables is reported in table 3 for each auditor class group for the matched-pair and switch samples. Some

---

[20] Inclusion of these explanatory variables, if relevant, will improve the specification of the regression equation. However, if these variables are measured with error and these errors are correlated with the earnings surprise proxy, the *ERC* estimates will remain biased. Nevertheless, so long as the degree of bias is the same for both auditor class groups, the test for a difference in the *ERC* between the two will remain informative.

[21] The statistics are reported for continuous measures of these control variables, not the binary measures used in the regression equations. They are calculated for observations used in the regressions, which exclude observations with missing data for some control variables (2 percent of matched-pair sample and 3 percent of switch sample).

[22] This is consistent with the results of Johnson and Lys (1990) and Teoh (1992b) that, on average, non-Big Eight to Big Eight auditor switches are preceded over a 60 day period by positive abnormal stock returns, and Big Eight to non-Big Eight switches by negative abnormal stock returns. This difference is captured by a significant intercept dummy in the estimated regression equation (6) reported in table 4.

Table 3
Correlation Matrix of Earnings Surprises and Control Variables
(Two-tailed p-values)

Panel A. Industry-Matched Sample:

	Audited by Big Eight Auditors						Audited by Non-Big Eight Auditors				
	UE	*MB*	*LMV*	β	*N*		*UE*	*MB*	*LMV*	β	*N*
UE	1.000	0.015	0.133	−0.062	0.097	*UE*	1.000	−0.046	0.031	−0.023	0.024
	0.000	0.585	0.000	0.026	0.000		0.000	0.108	0.275	0.423	0.392
MB		1.000	0.025	0.079	−0.009	*MB*		1.000	0.092	0.191	−0.001
		0.000	0.368	0.005	0.740			0.000	0.001	0.000	0.962
LMV			1.000	0.069	0.740	*LMV*			1.000	0.058	0.678
			0.000	0.014	0.000				0.000	0.042	0.000
β				1.000	0.083	β				1.000	0.046
				0.000	0.003					0.000	0.107
N					1.000	*N*					1.000
					0.000						0.000

Panel B. Switch Sample:

	Audited by Big Eight Auditors						Audited by Non-Big Eight Auditors				
	UE	*MB*	*LMV*	β	*N*		*UE*	*MB*	*LMV*	β	*N*
UE	1.000	0.024	−0.030	0.013	0.077	*UE*	1.000	0.102	0.122	−0.086	0.142
	0.000	0.695	0.624	0.834	0.212		0.000	0.094	0.044	0.160	0.019
MB		1.000	0.062	0.213	−0.020	*MB*		1.000	0.081	0.141	0.018
		0.000	0.315	0.000	0.741			0.000	0.184	0.020	0.767
LMV			1.000	0.152	0.625	*LMV*			1.000	0.042	0.669
			0.000	0.014	0.000				0.000	0.492	0.000
β				1.000	0.048	β				1.000	−0.009
				0.000	0.443					0.000	0.880
N					1.000	*N*					1.000
					0.000						0.000

Note: *UE*=unexpected earnings; *MB*=market value to book value; *LMV*=log of market value of common stock outstanding; β=market model slope coefficient; and *N*=number of analysts' forecasts in the consensus forecast.

differences in the partial correlations among the variables between the B8 and NB8 groups suggest that it might be appropriate to allow the slope coefficients of the control variables to vary between the two as in regression equation (6). The presence of significant correlations among some of the control variables suggests that it may be difficult to interpret the coefficients for the control variables separately even though they may be important joint determinants of the *ERC*.

V. Empirical Results

The OLS regressions for the matched-pair and switch samples are reported in table 4 and include the results when the market-adjusted return is used as a measure of

Table 4
OLS Regression of CARs on Earnings Surprises and Control Variables

$$CAR_{it}=\lambda_0+\lambda_1 D_{it}+\lambda_2 UE_{it}+\lambda_3 UE_{it}\bullet D_{it}+\lambda_4 UE_{it}\bullet MB_{it}+\lambda_5 UE_{it}\bullet MB_{it}\bullet D_{it}$$
$$+\lambda_6 UE_{it}\bullet\beta_{it}+\lambda_7 UE_{it}\bullet\beta_{it}\bullet D_{it}+\lambda_8 UE_{it}\bullet LMV_{it}+\lambda_9 UE_{it}\bullet LMV_{it}\bullet D_{it}$$
$$+\lambda_{10} UE_{it}\bullet\frac{1}{N_{it}}+\lambda_{11} UE_{it}\bullet\frac{1}{N_{it}}\bullet D_{it}+\epsilon_{it}$$

Panel A. Matched-Pair Sample:

CAR1 = cumulative market-adjusted return (F-statistic = 12.372, 0.0001; $\chi^2=47.359$, 0.0504; $\bar{R}^2=0.0465$)

Variable	*Parameter Estimate*	T *for* H0 (*parameter*=0)	*Probability* > |t|	*White's z-statistic*	*Probability* > |z|
Intercept	−0.015	−3.451	0.0006	−3.930	0.0001
D	0.004	0.589	0.5556	0.636	0.5246
UE	0.421	4.348	0.0001	3.804	0.0001
UE•D	−0.401	−2.523	0.0117	−1.972	0.0486
UE•MB	0.789	3.391	0.0007	2.732	0.0063
UE•MB•D	−0.405	−1.593	0.1113	−0.987	0.3235
UE•β	−0.183	−2.120	0.0341	−1.275	0.2022
UE•β•D	−0.029	−0.232	0.8162	−0.130	0.8960
UE•LMV	−0.229	−2.490	0.0128	−2.511	0.0120
UE•LMV•D	0.307	1.848	0.0648	1.783	0.0746
UE•$\frac{1}{N}$	0.060	0.689	0.4907	0.600	0.5488
UE•$\frac{1}{N}$•D	0.287	2.005	0.0450	1.505	0.1324

CAR2 = cumulative market model abnormal return (F-statistic = 3.813, 0.0001; $\chi^2=43.051$, 0.1131; $\bar{R}^2=0.0119$)

Variable	*Parameter Estimate*	T *for* H0 (*parameter*=0)	*Probability* > |t|
Intercept	−0.018	−3.669	0.0002
D	0.004	0.559	0.5765
UE	0.264	2.427	0.0153
UE•D	−0.317	−1.776	0.0759
UE•MB	0.428	1.637	0.1017
UE•MB•D	−0.102	−0.359	0.7197
UE•β	−0.101	−1.040	0.2985
UE•β•D	−0.134	−0.971	0.3317
UE•LMV	−0.212	−2.055	0.0400
UE•LMV•D	0.334	1.791	0.0734
UE•$\frac{1}{N}$	0.017	0.169	0.8657
UE•$\frac{1}{N}$•D	0.291	1.815	0.0697

Continued

abnormal stock return (*CAR*1) and those for the market model residual as the measure of abnormal returns (*CAR*2). The regression specification appears to be adequate in both panels A and B. The F-statistics for the goodness-of-fit are significant at the 0.0001 level in all four regressions. In addition, the χ^2-statistic to test for the presence of

Table 4–Continued

Panel B. Switch Sample:

*CAR*1 = cumulative market-adjusted return (F-statistic = 5.672, 0.0001; $\chi^2 = 28.724$, 0.6800; $\bar{R}^2 = 0.0869$)

Variable	*Parameter Estimate*	T *for* H0 (*parameter* = 0)	*Probability* > |t|
Intercept	0.014	1.581	0.1146
D	−0.033	−2.703	0.0071
UE	0.740	3.971	0.0001
UE•D	−0.974	−4.290	0.0001
UE•MB	−0.144	−0.503	0.6153
UE•MB•D	0.664	1.958	0.0508
UE•β	0.216	1.194	0.2330
UE•β•D	−0.560	−2.222	0.0267
UE•LMV	−0.606	−2.775	0.0057
UE•LMV•D	0.618	2.164	0.0309
$UE \cdot \frac{1}{N}$	−0.483	−2.569	0.0105
$UE \cdot \frac{1}{N} \cdot D$	1.032	3.953	0.0001

*CAR*2 = cumulative market model abnormal return (F-statistic = 4.374, 0.0001; $\chi^2 = 27.014$, 0.7590; $\bar{R}^2 = 0.0643$)

Variable	*Parameter Estimate*	T *for* H0 (*parameter* = 0)	*Probability* > |t|
Intercept	0.010	1.006	0.3151
D	−0.030	−2.172	0.0303
UE	0.857	3.986	0.0001
UE•D	−1.015	−3.874	0.0001
UE•MB	−0.430	−1.305	0.1926
UE•MB•D	0.788	2.012	0.0447
UE•β	0.078	0.371	0.7105
UE•β•D	−0.378	−1.299	0.1945
UE•LMV	−0.626	−2.487	0.0132
UE•LMV•D	0.521	1.581	0.1145
$UE \cdot \frac{1}{N}$	−0.439	−2.024	0.0434
$UE \cdot \frac{1}{N} \cdot D$	0.811	2.691	0.0074

Notes: D = dummy variable with a value of 1 for a Big Eight auditor, 0 otherwise; *UE* = unexpected earnings; *MB* = market value to book value; β = market model slope coefficient; *LMV* = log of market value of common stock outstanding; and N = number of analysts' forecasts in the consensus forecast. The model is from equation (6); MB_{it}, β_{it}, LMV_{it}, and $1/N_{it}$ are binary variables with a value of 1 if above median, and 0 otherwise. For the switch sample, the regression is estimated from two fiscal years before to two fiscal years after the switch.

heteroscedasticity of the residual errors is not significant in three of the four regressions. The exception is where the χ^2 = statistic for *CAR*1 is significant at the 5 percent level. Thus, we employ White's general heteroscedasticity-corrected covariance matrix estimate and report the resulting z-statistic and corresponding *p*-values.

The adjusted $\bar{R}^2$ in all four regressions is in the range reported by Lev (1989) in a survey of literature on earnings studies (usually less than 10 percent and generally between 1 percent and 5 percent in the majority of studies). In table 4, the adjusted $\bar{R}^2$ for the matched-pair sample is 4.65 and 1.19 percent, respectively, for the *CAR*1 and *CAR*2 regressions, and it is 8.69 and 6.43 percent, respectively, for the switch sample. The greater explanatory power for the switch sample than for the industry-matched sample suggests that using the firm's own past history as its control may better avoid missing determinants of the *ERC*. The explanatory power is higher for the *CAR*1 regression in both panels, which suggests, as noted earlier, that the market model may be a more noisy measure for expected returns.

In all four regressions, the evidence is consistent with the test hypothesis that the residual *ERC* is smaller for NB8 firms than for B8 firms. The λ_3 coefficient is negative, as predicted, and the magnitude is statistically significant at the 0.0001 level in both regressions of panel B, and at less than the 5 percent level (one-tailed) in panel A. The parameter estimate for λ_3 is -0.401 (White z-statistic $= -1.97$, $t = -2.53$) in panel A for the *CAR*1 regression and -0.317 ($t = -1.78$) for the *CAR*2 regression. In panel B, the λ_3 coefficient is -0.974 ($t = -4.29$) for the *CAR*1 regression and -1.015 ($t = 3.874$) for the *CAR*2 regression.

In all four regressions, the residual *ERC* for the B8 group is positive, ranging from 0.264 to 0.857, and is statistically significant at the 0.0001 level in three of the regressions and at 0.02 level in the fourth. In contrast, the residual *ERC* for the NB8 group (estimated by the sum of λ_2 and λ_3) is negative in three of the regressions and positive in the fourth. However, because the residual *ERC* measures only the part left unexplained by the included control variables, this does not mean that the *ERC* is positive for B8 groups and negative for NB8 groups.

Table 4 also provides some interesting evidence for comparison with previous studies on the determinants of the *ERC*. Because the qualitative results are similar for the *CAR*1 and *CAR*2 regressions, the remaining discussion comments only on *CAR*1 results. Consider first the results for the growth and persistence proxy. In panel A, the coefficient λ_4 for this proxy is positive and significant at the 1 percent level, as expected. The corresponding coefficient for the NB8 group is smaller, as evidenced by the negative λ_5, but also positive ($\lambda_4 + \lambda_5 > 0$). Thus, investors are more responsive to earnings surprises from firms that are growing and/or expected to persist.

In panel B of table 4, λ_4 is negative but not statistically significant, and λ_5 is positive and significant at the 5 percent level. They suggest that growth/persistence is relatively more important for the *ERC* of cross-class switches during the period when they are audited by NB8 auditors than when they are audited by B8 auditors. A partial explanation is that there is insufficient variation in the growth/persistence measure for these switches during the years when a B8 auditor is used and greater variation during the years when a NB8 auditor is used. For example, N-B switches might be experiencing high earnings growth just prior to the switch, and B-N switches might be experiencing declining earnings after the switch.

In panel A of table 4, the coefficient λ_6 for the risk proxy for the B8 group is negative though the White z-statistic is not statistically significant. The coefficient for the NB8 group is also negative and is not statistically different from the coefficient for the B8 clients. In panel B, λ_6 is positive though small and not statistically significant. The coefficient for the NB8 group is significantly more negative than for the B8 group, as evidenced by a statistically significant λ_7 coefficient. Thus, the risk proxy for the NB8 group in panel B is significantly negative. The generally weak evidence and inconsistent signs are not surprising, given the inconclusive results of previous studies and the possibility that β might proxy for both systematic risk of the firm and the amount of prior uncertainty of information, as discussed in section IV.

As mentioned earlier, the empirical evidence for the relation between client firm size and the *ERC* has been conflicting. This is not surprising given (1) the absence of a theory for the relevance of client firm size for the *ERC*, (2) the possibility that client firm size proxies for many relevant factors, and (3) the alternative methods for studying size in previous studies.[23] The results in our study are also conflicting. In both panels of table 4, the coefficient for the firm size proxy is significantly negative for B8 clients at the 1 percent level but is positive for NB8 clients. The difference between the coefficients for the groups is statistically significant.

Finally, no consistent results are obtained for 1/N, the reciprocal of the number of analysts following the firm. In panel A of table 4, the number of analysts appears unimportant for the *ERC*, as λ_{10} and λ_{11} are not statistically significant at the 5 percent level. In panel B, however, λ_{10} is significantly negative at the 1 percent level, $\lambda_{10}+\lambda_{11}$ is positive, and λ_{11} is significantly positive at the 1 percent level. This evidence suggests that the *ERC* increases (decreases) with the number of analysts for B8 (NB8) client firms. The model in section II predicts a positive relation between the *ERC* and the amount of prior uncertainty of the firm. If a small number of analysts following the firm implies that little is known about the underlying cash flows of the firm (greater prior uncertainty), then the model predicts a negative relation between the number of analysts and the *ERC*. Thus, the results for the NB8 group is consistent with the prediction of the model, but not those of the B8 group. Several explanations are possible for these weak and conflicting results. As seen in the correlation matrices (table 3), *N* is significantly correlated with firm size. The coefficients for 1/*N* and *LMV* therefore may be separately unreliable though both variables are important for the *ERC*. In addition, the relation between the amount of prior uncertainty for the firm and the number of analysts need not be monotonically decreasing (see O'Brien and Bhushan 1990). For example, greater uncertainty of the underlying cash flows might increase analysts' incentives to follow the firm because of the greater potential value of information.

VI. Conclusion

We have examined whether some auditors have higher quality than others, in the sense of generating earnings reports with greater credibility for investors. We tested this issue jointly with a prediction of a modified Holthausen and Verrecchia (1988) model, that earnings reports with less noise have higher *ERC*s. The empirical test is operationalized as a test for a difference in the *ERC* between B8 and NB8 clients in a regression of abnormal stock returns on earnings surprise and control variables for other determinants of the *ERC*. We find that B8 clients have statistically significantly larger *ERC*s than NB8 clients. The result obtains in both a matched-pair sample of firms paired according to industry membership and a switch sample of firms grouped according to auditor size. Thus, the evidence is consistent with the joint hypothesis that larger auditors generate more precise earnings and that the Holthausen-Verrecchia prediction relating *ERC* to earnings precision is correct.

[23] See Shevlin and Shores (1990, fn. 21) for a comprehensive summary of the treatment and findings for firm size in past earnings studies.

The study also provides new evidence regarding the determinants of the ERC. Previous studies have found that the ERC increases with growth and persistence, have obtained inconsistent findings for client firm size and risk across studies, and have not considered the number of analysts following a firm as a determinant. We find that the relation of the ERC with growth and persistence is positive, and with firm risk is weak and negative. The relation of the ERC with the number of analysts following the firm, as a proxy for the noise in the information environment, is found to differ between B8 and NB8 client firms; the relation is positive for B8 clients, negative for NB8 clients. The relation of the ERC with client firm size is also found to differ between the groups; it is negative for B8 clients and positive for non-Big Eight clients.

We provide robustness checks on the results by using two alternative sample designs and including known determinants of the ERC as control variables. However, the results may be biased if there are other determinants of the ERC, not fully captured by the included control variables, that vary systematically between the auditor class groups. In addition, measurement error of the included variables that differ between the groups will also bias the results. However, to the extent that the included control variables are related to the ERC by virtue of their correlation with auditor size and credibility, the test procedure provides a conservative estimate of the effect of auditor size on the ERC.

By linking differences in perceived auditor quality to the ERC, this study provides evidence for the view that larger auditors are more credible. The issue of whether auditor quality is related to size may be relevant for the current policy debate on whether to permit mergers among large audit firms. This study is also relevant for capital markets research because it provides evidence of additional determinants of the ERC, auditor size, and the number of analysts following the firm. Finally, the study provides the first empirical verification of the Holthausen-Verrecchia model of earnings precision and stock price changes.

References

Abdel-khalik, A. R., and I. Solomon, eds., 1988. *Research Opportunities in Auditing: The Second Decade*. Sarasota, FL: AAA Auditing Section.

Atiase, R. 1985. Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research* 23 (Spring): 21–35.

Beatty, R. P. 1989. Auditor reputation and the pricing of initial public offerings. *The Accounting Review* 64 (October): 693–709.

Biddle, G., and G. Seow. 1991. The estimation and determinants of associations between returns and earnings: Evidence from cross-industry comparisons. *Journal of Accounting, Auditing & Finance* 6 (Spring): 183–232.

Brown, P., G. Foster, and E. Noreen. 1985. *Security Analyst Multi-year Earnings Forecasts and the Capital Market*. Sarasota, FL: AAA.

Carpenter, C., and R. Strawser. 1971. Displacement of auditors when clients go public. *Journal of Accountancy* 131 (June): 55–58.

Christie, A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting & Economics* 9 (December): 231–58.

Collins, D. W., and S. P. Kothari. 1989. An analysis of intertemporal and cross-sectional determinants of earnings response coefficient. *Journal of Accounting & Economics* 11 (July): 143–81.

Datar, S., G. Feltham, and J. Hughes. 1991. The role of audits and audit quality in valuing new issues. *Journal of Accounting & Economics* 14 (March): 3–49.

DeAngelo, L. 1981. Auditor size and audit quality. *Journal of Accounting & Economics* 3 (December): 183–99.

Dopuch, N., and D. Simunic. 1980. The nature of competition in the auditing profession: A descriptive and normative view. In *Regulation and the Accounting Profession*, edited by J. Buckley and F. Weston. Belmont, CA: Lifetime Learning Publication.

——, and ——. 1982. Competition in auditing: An assessment. *Fourth Symposium on Auditing Research.* Urbana, IL: Office of Accounting Research: 401–05.

Easton, P., and M. Zmijewski. 1989. Cross-sectional variation in the stock market response to accounting earnings announcements. *Journal of Accounting & Economics* 11 (July): 117–42.

Holthausen, R., and R. Verrecchia. 1988. The effects of sequential information releases on the variance of price changes in an intertemporal multi-asset market. *Journal of Accounting Research* 26 (Spring): 82–106.

Imhoff, E. A. 1988. A comparison of analysts' accounting quality judgments among CPA firms' clients. *Auditing: A Journal of Practice & Theory* 7 (Spring): 182–91.

——, and G. J. Lobo. 1992. The impact of *ex ante* earnings uncertainty on earnings response coefficient. *The Accounting Review* 67 (April): 427–39.

John, T. 1991. Optimal size and structure of auditing firms. Working paper, New York University, New York.

Johnson, B., and T. Lys. 1990. The market for audit services: Evidence from voluntary auditor changes. *Journal of Accounting & Economics* 12 (January): 281–308.

Kormendi, R., and R. Lipe. 1987. Earnings innovations, earnings persistence and stock returns. *Journal of Business* 60 (July): 323–45.

Lang, M., and M. McNichols. 1990. Earnings quality, financial distress and the incremental information content of cash flows. Working paper, Stanford University, Palo Alto.

Leamer, E. 1978. *Specification Searches: Ad Hoc Inference with Nonexperimental Data.* New York: John Wiley and Sons.

Lev, B. 1989. On the usefulness of earnings and earnings research: Lessons and directions from two decades of empirical research. *Journal of Accounting Research* 27 (Supplement): 153–92.

Lipe, R. 1990. The relation between stock returns and accounting earnings given alternative information. *The Accounting Review* 65 (January): 49–71.

Marais, L., K. Schipper, and A. Smith. 1989. Wealth effects of going private for senior securities. *Journal of Financial Economics* 23 (June): 155–91.

Nichols, D., and D. Smith. 1983. Auditor credibility and auditor changes. *Journal of Accounting Research* 21 (Autumn): 534–44.

O'Brien, P. 1988. Analysts' forecast as earnings expectations. *Journal of Accounting & Economics* 10 (January): 53–83.

——, and R. Bhushan. 1990. Analyst following and institutional ownership. *Journal of Accounting Research* 28 (Supplement): 55–76.

O'Keefe, T., and R. Barefield. 1985. Purchase of accounting services by small clients: Description and analysis. Working paper, University of Oregon, Eugene.

Palmrose, Z. 1988. An analysis of auditor litigation and audit service quality. *The Accounting Review* 63 (January): 55–73.

Shevlin, T., and D. Shores. 1990. Firm size, security returns, and unexpected earnings. Working paper, University of Washington, Seattle.

Simunic, D., and M. Stein. 1987. Product differentiation in auditing: Auditor choice in the market for unseasoned new issues. Research Monograph 13, *The Canadian Certified General Accountants' Research Foundation*, Vancouver, B.C.

Teets, W. 1992. An empirical investigation into the association between stock market responses to earnings announcements of electric utilities. *Journal of Accounting Research* 30 (Autumn): 274–85.

Teoh, S. H. 1992a. Auditor independence, dismissal threats, and the market reaction to auditor switches. *Journal of Accounting Research* 30 (Spring): 1–23.

——. 1992b. The effect of auditor switches on accounting and market measures of value. Working paper 89-4, Anderson Graduate School of Management, University of California–Los Angeles.

Titman, S., and B. Trueman. 1986. Information quality and the valuation of new issues. *Journal of Accounting & Economics* 8 (June): 159–72.

SHORTER ARTICLES

1. Li-Chin Jennifer Ho
   Option Trading and the Relation Between Price and Earnings: A Cross-Sectional Analysis
2. Norman Strong and Martin Walker
   The Explanatory Power of Earnings for Stock Returns
3. Alan S. Dunk
   The Effect of Budget Emphasis and Information Asymmetry on the Relation Between Budgetary Participation and Slack

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 368–384

Option Trading and the Relation Between Price and Earnings: A Cross-Sectional Analysis

Li-Chin Jennifer Ho
University of Texas at Arlington

SYNOPSIS AND INTRODUCTION: Prior research suggests that option trading affects the availability and timeliness of predisclosure information about firms and that the price-earnings relation is influenced by characteristics of the predisclosure information environment. Motivated by these research findings, this study (1) examines various firm-specific attributes that are likely to explain the different information environments of firms with and without exchange-traded options, and (2) investigates the price-earnings relation of such firms. The price-earnings relation is examined in both "event study" (short-window) and "association study" (long-window) contexts. The short-window analysis tests the hypothesis that the "surprise" in quarterly earnings reports (measured by the abnormal stock return variability around earnings announcements) is greater for nonoption firms than for option firms (hypoth. H1). The long-window analysis tests the hypothesis that security prices signal future earnings changes earlier for option firms than for nonoption firms (hypoth. H2A). The hypothesis also predicts that nonoption firms are more likely than option firms to exhibit post-FYE (fiscal year-end) drift (hypoth. H2B).

The results indicate that option firms are associated with five firm-specific attributes: (1) larger firm size, (2) higher institutional concentration, (3) higher analyst coverage, (4) higher trading volume, and (5) more *Wall Street Journal Index* news releases. Based on 3,721 quarterly earnings announcements of 431 firms during the 1980–1983 period, the results support the first hypothesis; abnormal return variability surrounding quarterly earn-

This article is based in part on my dissertation, completed at the University of Texas, Austin. I am grateful to my dissertation committee, Robert Freeman, Ross Jennings, Thomas Shively, Laura Starks, and especially Robert May (chairman). I gratefully acknowledge the valuable comments of the Associate Editor and two anonymous referees. I appreciate John Hassell and Jeffrey Tsay for many useful discussions. I gratefully acknowledge the contribution of I/B/E/S Inc. for providing earnings per share forecast data, available through the Institutional Brokers Estimate System. Financial support from the Department of Accounting at the University of Texas at Arlington is gratefully acknowledged. I alone am responsible for any errors or omissions.

Submitted April 1991.
Accepted December 1992.

ings announcements is significantly greater for nonoption firms than for option firms. To examine how the five firm-specific attributes provide alternative explanations for the observed results, the empirical tests are repeated by using a unilateral matching approach. Five subsamples (each based on one of the five variables) are constructed. The results for each support the first hypothesis.

In addition, a control portfolio is constructed that conservatively controls for the five proxy variables (i.e., the *nonoption* firms in this portfolio are associated with *larger* firm size, *higher* institutional concentration, *higher* analyst coverage, *higher* trading volume, and *more Wall Street Journal Index* news releases). With a relatively smaller sample size, the control portfolio results also support the first hypothesis. Finally, multiple regression results indicate that option trading possesses incremental explanatory power over the other variables in explaining the differential content of earnings releases for option and nonoption firms. The only other variable that is significant in the regression analysis is the number of *Wall Street Journal Index* news releases, which is negatively related to measured information content. This suggests that press coverage is also an important variable in explaining cross-sectional differences in information content, at least for the sample firms.

With respect to the second hypothesis, the results for the entire sample, the matched subsamples, and the control portfolio indicate that the security prices of option firms anticipate accounting earnings earlier than do those of nonoption firms. For option firms, about 50 percent of the price change associated with economic events contributing to the current year's earnings change is realized in the previous fiscal year. For nonoption firms, however, a significant portion (about 70 percent) of the price change associated with the earnings change occurs in the current fiscal period and in the 12 months following the fiscal year-end. In addition, the magnitude of the post-FYE drift is greater for nonoption firms than for option firms.

Certain important *caveats* apply to the results. First, given that the tests for the matched subsamples attempt to control only for one factor at a time, the results are subject to a missing-variable problem. Second, while the option-trading variable exhibits statistical significance in the regression analysis, the explanatory power of the full regression is low (although it is similar to that achieved in typical cross-sectional analyses of abnormal security returns). Finally, given the lack of a formal theory concerning option trading and information flows, it is difficult to infer that option trading *causes* the differences in the price-earnings relation across firms.

Key Words: *Option trading, Informativeness of security prices, Relative stock return variability, Cumulative abnormal returns.*

Data Availability: *The analysts' earnings forecast data from the Institutional Brokers Estimate System (I/B/E/S) are proprietary and can be obtained from I/B/E/S Inc. The remainder of the data can be gathered from sources identified in the text.*

THE remainder of the article is organized as follows. I review the relevant literature and discuss how this study differs from others in section I, then I discuss the methodology, describe the sample, and report the empirical results for the short-window analysis in section II. The methodology, sample, and evidence for the long-window analysis are presented in section III, and concluding remarks are offered in section IV.

I. Prior Literature and Hypotheses

The primary motivation for this study stems from findings in previous research that option listing and subsequent trading increase available information. Prior work (see, e.g., Black 1975; Cox and Rubinstein 1985; Diamond and Verrecchia 1987; Ross 1976) suggests that option trading increases investors' incentives to collect private information because options allow investors to use that information in ways that are not available by trading in stocks alone. Also, options markets may increase the speed at which private information is impounded into security prices (Diamond and Verrecchia 1987). Empirical research on the interaction between stock and option markets provides evidence consistent with these arguments (see, e.g., Chiras and Manaster 1978; Jennings and Starks 1986; Manaster and Rendleman 1982; Patell and Wolfson 1979, 1981; Skinner 1990).

This study is also motivated by the empirical literature on characteristics of information environments that affect cross-sectional differences in the security price reaction to accounting earnings releases. This line of research suggests that the price-earnings relation is influenced by proxies for the level of interim information, such as firm size, analyst following, and the degree of financial press coverage. Representative studies include Grant (1980), Atiase (1985, 1987), Freeman (1987), Collins et al. (1987), Dempsey (1989), Lobo and Mahmoud (1989), and Shores (1990).

The option-trading arguments coupled with the findings of studies of the information environment suggest that option trading may affect the relation between security returns and accounting earnings. In general, the price-earnings relation is examined in either an "event study" or an "association study" context. Event studies focus on whether the earnings announcement contains some surprise element that causes investors to revise their expectations of future cash flows (see, e.g., Beaver 1968). Association studies concern the comovement between security prices and valuation-relevant economic events that are captured by the earnings determination process (see, e.g., Ball and Brown 1968). This study examines option trading and the price-earnings relation by using both methods.

First, option trading is expected to reduce the surprise in accounting earnings. Prior studies of information content indicate that stock return variability tends to increase around announcements of accounting earnings (see, e.g., Beaver 1968). This increase is considered evidence that the information in the disclosures causes revision of beliefs. Given the argument that option trading increases predisclosure information, one would expect earnings announcements of nonoption firms to be relatively more informative for investors and to cause greater belief revision, which in turn should generate higher return variability. This suggests the first hypothesis in this study:

H1: The surprise associated with quarterly earnings announcements (as proxied by the abnormal stock return variability surrounding earnings announcements) is greater for nonoption firms than for option firms.

Second, option trading is expected to improve the extent to which price changes anticipate future accounting earnings changes. The lead-lag relation between changes in price and earnings is documented by Ball and Brown (1968) and Beaver et al. (1980). Recent research by Freeman (1987) and Collins et al. (1987) further suggests that the degree of anticipation is positively related to firm size, which is a proxy for differences in information environment. In particular, they find that a significant portion of the total price change associated with the current year's earnings change is realized in the previous fiscal year for large, rather than small, firms. They also document that the cumulative abnormal returns (CARs) of large firms exhibit an earlier post-FYE decline than do those of small firms. Since option trading is argued to be associated with a broader and richer information set about a firm's activities, one would expect that (1) security prices signal future earnings changes earlier for option firms than for nonoption firms and (2) nonoption firms are more likely to exhibit post-FYE drift. This motivates the second hypothesis in this study.

H2A: The proportion of CARs associated with the current year's earnings changes that occurs in the previous fiscal year (more generally, at earlier points in time) is greater for option firms than for nonoption firms.

H2B: The magnitude of post-FYE drift is greater for nonoption firms than for option firms.

This study essentially extends and complements work by Skinner (1990) and Jennings and Starks (1986). Skinner examines the stock price reaction to earnings announcemnts around the time of option listing. Using two information content measures (the earnings response coefficient and the size of the abnormal returns around earnings announcement dates), Skinner finds that the information content of firms' earnings releases is smaller after their stocks are listed on option exchanges.

In contrast to Skinner's (1990) examination of how information content varies over time, this study compares price reactions across firms. One difficulty with Skinner's longitudinal test is that his analysis is unable to control for the statistically significant differences in firm size for the periods before and after listing. He finds that the size of option firms increases significantly after listing, which could partially explain the decline in information content (see his table 5). This study strengthens Skinner's findings by using a cross-sectional design that specifically controls for firm size (and other known variables that may provide alternative explanations; see section II).

Although Skinner (1990, table 4) documents the smaller price reaction to quarterly earnings announcements, he fails to find a significantly positive association between option listing and preemption of the information content of quarterly earnings prior to announcement. This result is inconsistent with the option-trading argument and prior studies of information environment that suggest a positive association between the extent of the preemption and the various proxies for the level of interim information (see, e.g., Collins et al. 1987; Freeman 1987; Shores 1990). Given this puzzling finding, this study provides additional evidence on the preemptive role of option trading by using cross-sectional data and different cumulation periods. In particular, this study involves a test to examine how option trading affects the extent to which price changes lead earnings changes by using cumulation periods similar to those in Collins et al. (1987) and Freeman (1987).

Jennings and Starks (1986) document that the stock prices of nonoption firms take longer to adjust to accounting earnings releases than do those of option firms. This

Size, Beta, and Trading Volume. As discussed in Skinner (1990), the optio changes have incentives to list options on firms with a high level of investor inte trading activity. Size and riskiness are important variables because investor i generally increases with the size of the firm, and option value increases with the variance (and so the general riskiness) of the stock. Trading volume of common is also a key indicator of investor interest. Stocks with listed options are likel traded most actively.

Comparisons of size and trading volume for option and nonoption fir important because previous research suggests that the price-earnings relation de on these two variables. Atiase (1985), Collins et al. (1987), and Freeman (1987) in that firm size is positively associated with the amount of predisclosure infor and, therefore, is negatively related to the information content of accounting ea releases. Shores (1990) documents evidence consistent with the notion that thinness (measured by trading volume as a percentage of outstanding sha negatively related to the level of interim information.

As shown in table 1, the median market value of the option firms is about twi of the nonoption firms, and the mean value is about three times larger. differences are statistically significant at the 0.01 level with a Wilcoxon rank-sum As also shown, over the period 1980–1983, the median size of firm-specific syst risk coefficients (estimated by conventional market model regressions with returns data for the relevant calendar years) is significantly greater for the option than for the nonoption firms at the 0.01 level (1.125 vs. 0.669).

Consistent with prior research, the results indicate that the option firms are more actively traded in the stock market, as evidenced by a significantly monthly trading volume. The median (mean) monthly volume for the option fi 2.50 (2.89) million shares. In contrast, the median (mean) monthly volume f nonoption firms is 0.77 (0.94) million shares. These differences are significant 0.01 level.[2]

Earnings Volatility. Earnings volatility matters in this context because (1983) documents that the magnitude of stock price reaction to ear announcements (measured by the announcement period return variances) is pos associated with the variability of earnings changes. The results in table 1 indica the median and mean standard deviations of annual earnings-per-share ch (estimated from the time series of annual earnings prior to the year the firms ap in the sample) are similar for both groups. The median magnitude of the sta deviation is 1.215 for the option firms and 1.185 for the nonoption firms. differences are associated with a two-tailed Wilcoxon probability of 0.235. Thus, and nonoption firms in this sample do not differ systematically in terms of ea volatility.

Analyst Following, Institutional Holdings, and Financial Press Coverage. Res by Grant (1980), O'Brien and Bhushan (1990), and Shores (1990) suggests th amount of information publicly available about a firm is positively associated w the number of financial analysts following the firm, (2) the number of institutional holders of the firm's stocks, and (3) the number of news announcements in *The Wall Street Journal Index* (as a proxy for the extent to which a firm is followed by the financial press).

---

[2] In conformance with Shores (1990), this study also compares the distribution of "percentage volume" across firms. For each firm in each year, percentage trading volume is calculated as the firm's monthly volume divided by the number of common shares outstanding as of the last trading day of th year. The analysis based on percentage trading volume also indicates that the option firms are muc actively traded in the stock market.

The results in table 1 show that option firms are more closely followed by analysts. The median number of analysts (as proxied by the number reported in the I/B/E/S Summary History tape) is 11.53 for the option firms and 8.88 for the nonoption firms. Moreover, the option firms have more institutional investors (a median of 189 institutions vs. 120). Finally, the option firms have more extensive coverage in *The Wall Street Journal Index*. The median number of news announcements for the option firms is 28.33, versus 20.50 for the nonoption firms. All these differences are significant at the 0.01 level with a two-tailed Wilcoxon test.

In summary, option and nonoption firms differ with respect to their systematic risk and five other variables associated with firms' information environments. The comparisons of systematic risk indicate that risk-adjusted abnormal returns, rather than raw returns, are a more appropriate dependent variable for the hypotheses in this study. With respect to the remaining five variables, previous research suggests that they may provide alternative explanations for any statistically significant results; therefore, they must be considered in both the short- and long-window analyses.

Methodology

The statistical methodology used in the short-window analysis is from Patell (1976) and Dodd et al. (1984). First, the market model is used to adjust for the effect of contemporaneous market-wide information releases. The prediction error, u_{it}, during the announcement period (AP) is obtained from the market model:

$$u_{it}=R_{it}-(\hat{a}_i+\hat{b}_i R_{mt}) \qquad t\in AP,$$

where R_{it} is the actual return on the common stock of firm i on day t; R_{mt} is the return on the CRSP value-weighted market portfolio on day t; $\hat{a}_i$ and b_i are market model coefficients estimated over a 200-day trading period from day −209 to day −10, where day 0 is the earnings announcement date. The announcement period comprises days −1, 0, and +1 relative to the earnings announcement date.

In conformance with Patell (1976) and Dodd et al. (1984), for each firm i and the three-day period, a statistic, RRV_i, is calculated,

$$RRV_i=\frac{\sum_{t=-1}^{1}\frac{u_{it}^2}{c_{it}\times s_i^2}}{\sqrt{3}}*[(T_i-4)/(T_i-2)],$$

where T_i is the number of days in the estimation period for firm i, s_i^2 is the variance of the residuals during the estimation period for firm i, and c_{it} is an adjustment factor in variance because of prediction outside the estimation period for firm i on day t. The abnormal return variability for firm i during the earnings announcement period is measured by RRV_i. For an intuitive interpretation, it is a ratio of the new information conveyed by the earnings announcement to the average information level during the nonannouncement period.

The short-window analysis treats each firm as a single observation and calculates a mean *RRV*-statistic for each firm over the period 1980–1983. The distribution of these firm-specific means is then compared for the option and nonoption firms. The statistical significance of empirical results is assessed by using two nonparametric tests: a two-sample Wilcoxon (rank sums) test and a nonparametric test based on the empirical distribution of firm-specific mean *RRV*-statistics.[3]

Results

Entire Sample (no control for self-selection bias). The RRV-statistics for the entire sample of option and nonoption firms over the years 1980–1983 are reported first in table 2. The median firm-specific *RRV*-statistic for the option firms is 2.011, versus 2.258 for the nonoption firms. The difference is associated with a one-tailed Wilcoxon probability of 0.009. The mean firm-specific *RRV*-statistic for the nonoption firms exceeds that for the option firms by 0.165, a difference that is significant at the 0.09 level when using the empirical distribution generated from a randomization test (see fn. 3). Thus, the evidence is consistent with the prediction that the surprise in quarterly earnings announcements is greater for nonoption firms than for option firms.[4]

Unilateral Matching. To examine how the five firm-specific characteristics provide alternative explanations for the results of the entire sample, a unilateral matching approach is employed. Five matched subsamples (each based on one of the five attributes) are constructed by pooling all the firms and ranking them on the relevant attribute (e.g., their average stock market values over the years 1980–1983). A firm is discarded if the observations in both the next highest and next lowest rank are of the same type (option or nonoption). In each subsample, both Wilcoxon and two-sample *t*-tests fail to detect any significant differences in the matching variable between the matched option and nonoption firms.

Note that the test for each matched subsample attempts to control for only one factor at a time.[5] Thus, it is difficult to establish that option listing *per se* (i.e., after con-

---

[3] Marais (1984) indicates that the conventional significance tests of Patell *U*-statistics, which assume normality of market model prediction errors, are biased toward rejection of the null hypothesis of no information content. Given this concern, this study reports only the two nonparametric test results.

In conformance with Noreen (1989), a randomization procedure is employed to generate the empirical distributions of firm-specific mean *RRV*-statistics as follows: (1) Randomly shuffle the sample firms and form two pseudo portfolios: option and nonoption. (2) Calculate the mean *RRV*-statistic for each firm over the sample period and then compute the average of these firm-specific means for each pseudo portfolio. Let *ARRVO* (*ARRVN*) = the average size of firm-specific mean *RRV*-statistics for the option (nonoption) pseudo portfolio. Define *DARRV*=*ARRVN*−*ARRVO*. (3) Repeat the previous steps 1,000 times and use the distributions of the 1,000 values of *DARRV* to assess the significance level of the actual realization of *DARRV*.

[4] Four earnings announcements with extreme *RRV*-statistics (94.45, 81.33, 49.99, and 46.53) were deleted. The results prior to this deletion are slightly less significant but qualitatively similar to those reported in table 2. To examine the extent to which outliers may affect the results, all observations greater than 30 and then all greater than 20 were deleted in turn. The overall results are not materially affected after these increasingly stringent outlier deletion rules are employed.

[5] Although the two groups in the size subsample appear similar in analyst following and institutional holdings, the analysis indicates that the size-matched option firms are more closely followed by *The Wall Street Journal Index*. The median number of news releases for the option firms is 24, relative to 21 for the nonoption firms (the difference is significant at the 0.01 level with a Wilcoxon test). In addition, the option firms are significantly more actively traded; the median monthly trading volume for the option firms is 1.82 million shares, compared to 0.85 million for the nonoption firms ($p < 0.001$ with a Wilcoxon test). Finally, the analysis indicates that (1) the analyst subsample does not control for the other four variables; (2) the institutional holding subsample controls only for institutional holdings, size, and analyst following; (3) the financial news subsample controls only for news coverage and analyst following; and (4) the trading volume subsample controls for trading volume but not other variables.

Table 2
Comparison of Firm-Specific Mean *RRV*-Statistics
Calculated over Three-Day Announcement Period
(1980–1983)

	Number of Firms	*Mean*	*Median*
Entire Sample			
Option	255	2.282	2.011
Nonoption	176	2.447	2.258
p-value for difference		0.090	0.009
Size Subsample			
Option	118	2.190	1.818
Nonoption	118	2.477	2.208
p-value for difference		0.060	0.009
Analyst Subsample			
Option	132	2.218	1.982
Nonoption	132	2.468	2.216
p-value for difference		0.070	0.008
Institutional Holding Subsample			
Option	121	2.125	1.949
Nonoption	121	2.440	2.256
p-value for difference		0.040	0.004
Financial News Subsample			
Option	122	2.251	2.014
Nonoption	122	2.437	2.205
p-value for difference		0.130	0.070
Trading Volume Subsample			
Option	80	2.187	1.763
Nonoption	80	2.413	2.244
p-value for difference		0.100	0.010
Control Portfolio			
Option	51	2.251	1.661
Nonoption	51	2.640	2.277
p-value for difference		0.030	0.010

Note: The *RRV*-statistics are calculated by using a three-day earnings announcement period (days −1, 0, and +1). Day 0 is defined as the earlier of the declaration date of the quarterly earnings announcement on the COMPUSTAT quarterly industrial file or the announcement date in *The Wall Street Journal Index*. The market model is estimated over a 200-day trading period from day −209 to −10. Those earnings announcements with *RRV*-statistics greater than 40 are deleted because these extreme observations are considered to be outliers (see fn. 4). The *p*-value for mean is the probability level from the empirical distributions generated by a randomization test (see fn. 3); the *p*-value for median is the one-tailed probability level from a two-sample Wilcoxon test of the alternative hypothesis that the distribution of firm-specific *RRV*-statistics is greater for nonoption than for option firms.

trolling for all known variables simultaneously) has an effect on the price-earnings relation because the tests are subject to a missing-variable problem. With this limitation in mind, the results indicate that, in each subsample, the stock price reaction to quarterly earnings announcements is greater for the nonoption firms than for the matched option firms. The Wilcoxon two-sample test supports hypothesis H1 for all subsamples: the median *RRV*-statistic is significantly greater for the nonoption firms than for the option firms at the 0.01 level for size, analyst, institutional holding, and trading volume subsamples. The difference is weakly significant ($p=0.07$) for the financial news subsample. The randomization test also rejects the null hypothesis of no differences in mean *RRV*-statistic at the 0.10 level for all subsamples except that for *The Wall Street Journal* coverage.

Control Portfolio. To mitigate concern about a missing-variable problem, an alternative control portfolio is formed out of the size subsample so each option firm would have both lower trading volume and fewer news releases than its matching nonoption firm. If these two factors are inversely related to the information content of accounting earnings announcements (as suggested by prior research), then one would expect higher information content for the earnings releases of the option firms. If their earnings releases have lower information content, the evidence is consistent with the price-earnings relation's being affected by option listing.

The above procedure produces a 102-firm portfolio (51 option and 51 nonoption firms) in which both the trading volume and number of *The Wall Street Journal Index* releases are significantly higher for the nonoption firms than for the option firms at the 0.01 level. Further examination reveals that the nonoption firms in this control portfolio are associated with larger firm size, higher institutional concentration, and higher analyst coverage (all significant at the 0.01 level). Thus, the control portfolio also conservatively controls for the effects of the remaining variables. Given extant evidence that these five factors are inversely related to the information content of earnings releases, a significantly higher *RRV*-statistic for the nonoption firms in this control portfolio would therefore be attributable to option listing, not to the effects of the five proxy variables.

The results indicate that the median (mean) *RRV*-statistic for the option firms is 1.661 (2.251) versus 2.277 (2.640) for the nonoption firms. The difference is significant at the 0.01 (0.03) level with the Wilcoxon test (randomization test). Consequently, even though the control portfolio analysis is conservatively biased against finding significance, the results suggest that accounting earnings releases are more informative for nonoption firms than for option firms.[6]

Multiple Regression Analysis. Table 3 presents the results of regressing firm-specific *RRV*-statistics on the following variables: stock market value, number of institutional holders, number of financial analysts, number of *The Wall Street Journal Index* entries, monthly trading volume, and a dummy variable for option-listing status (coded zero for nonoption firms and one for option firms). The multiple regression analysis provides evidence on the incremental explanatory power of option trading over the other variables.

As shown, three of the variables (i.e., the option-trading dummy variable, the number of news releases, and monthly trading volume) exhibit coefficients consistent with the predicted sign. The coefficients on the dummy variable and number of news releases are significant at the 0.04 level (one-tailed), which indicates that option listing and press coverage are significant in explaining cross-sectional variation in security price reaction to quarterly earnings releases after controlling for the effects of other

---

[6] To shed more light on the missing-variable problem, the intersection of the various matched subsamples across the five criteria is examined. Fourteen option and 28 nonoption firms are included in the intersection subsample, and the analysis indicates that this sample also conservatively controls for the effects of the five proxy variables. The *RRV* results also support an option-trading effect.

Table 3

Regressions of the RRV-Statistics on Measures of Option-Trading Status, Size, Institutional Holdings, Analyst Following, News Releases, and Trading Volume

(n=431 firms; t-statistics in parentheses)

$$Ln(RRV_i)=\alpha+\beta_1{*}D_i+\beta_2{*}Ln(SIZE_i)+\beta_3{*}Ln(INST_i)+\beta_4{*}FINA_i+\beta_5{*}WSJ_i+\beta_6{*}Ln(TRAD_i)+\epsilon_i$$

α	β_1	β_2	β_3	β_4	β_5	β_6
0.670	−0.124	0.020	0.064	0.009	−0.002	−0.046
(0.714)	(−1.735)	(0.383)	(0.858)	(1.091)	(−1.776)	(−0.966)

F=2.914 (p=0.009) Adjusted R^2=0.026

Note: D_i=dummy variable indicating option-trading status (coded 0 for nonoption firms and 1 for option firms);
$SIZE_i$=stock market value for firm *i*;
$INST_i$=number of institutional holders for firm *i*;
$FINA_i$=number of financial analysts following firm *i*;
WSJ_i=number of news articles in *The Wall Street Journal Index* for firm *i*; and
$TRAD_i$=monthly trading volume for firm *i*.

known variables.[7] However, collinearity diagnostics indicate high correlations among the independent variables, and the coefficient of determination for the regression equation in this analysis is as low as 0.026 (although it is similar to that achieved in typical cross-sectional analyses of abnormal security returns). Therefore, the results must be interpreted with caution.

III. The Long-Window Analysis

Methodology

The long-window tests concern the pattern of return changes relative to earnings changes for option versus nonoption firms. In conformance with Freeman (1987) and Collins et al. (1987), the following empirical procedures are constructed to test hypotheses H2A and H2B. First, all available firms are divided into good and bad news portfolios. The classifications for each sample year (1980–1983) are based on the direction of the change in income before extraordinary items and discontinued operations ($DIBED \equiv IBED_y - IBED_y - 1$). A firm is placed in a good news portfolio if its *DIBED* is positive and in a bad news portfolio if *DIBED* is negative. Income information is from the COMPUSTAT annual industrial file.

Second, abnormal returns for each firm are calculated over the 36-month interval beginning 24 months prior to the end of the relevant fiscal year and ending 12 months after that fiscal year-end. Abnormal returns are obtained from the market model by using the CRSP returns file. The market model coefficients are estimated by OLS regressions over a 60-month period preceding the accumulation of abnormal returns.

---

[7] Because the distribution of the RRV-statistics is skewed, the regressions reported in table 3 use the natural logarithm of the RRV-statistics as the dependent variable. Results are similar, although less significant, when the RRV-statistics are used as the dependent variable.

Third, separate portfolios of option (O) and nonoption (N) firms are constructed of long positions in good news firms and short positions in bad news firms. The following composite $CCAR_t$ score is calculated for both option and nonoption portfolios,

$$CCAR(O)_t = 1/N_o\left(\sum_{i=1}^{G_o} CAR_{it} - \sum_{i=1}^{B_o} CAR_{it}\right),$$

$$CCAR(N)_t = 1/N_n\left(\sum_{i=1}^{G_n} CAR_{it} - \sum_{i=1}^{B_n} CAR_{it}\right),$$

where CAR_{it} is the cumulative abnormal returns of firm i for month t ($t=1,\ldots,36$); G_o (B_o) is the number of good (bad) news firms in the option portfolio; G_n (B_n) is the number of good (bad) news firms in the nonoption portfolio; N_o (N_n) is the total number of firms in the option (nonoption) portfolio. Therefore, $N_o = G_o + B_o$ and $N_n = G_n + B_n$.

The long-window hypothesis is tested by comparing the percentage of $CCAR(O)$ versus the percentage of $CCAR(N)$ explained in different points of time. Let $CCAR(O)_{36}$ represent option-firm CCAR in month +36 (i.e., 12 months after the relevant fiscal year-end), and let $O_t = CCAR(O)_t / CCAR(O)_{36}$. Similarly, let $N_t = CCAR(N)_t / CCAR(N)_{36}$. Thus, O_{12} represents the percentage of option firms' 36-month composite CCAR that occurs at the end of the previous fiscal year (month 12). And N_{12} represents the percentage of nonoption firms' 36-month composite CCAR that is realized at the end of the previous fiscal year.

Sample

The initial sample in the long-window test comprises those firms in the previous short-window test. To be included in the final sample of a particular sample year for the long-window test, a firm must meet two additional criteria:

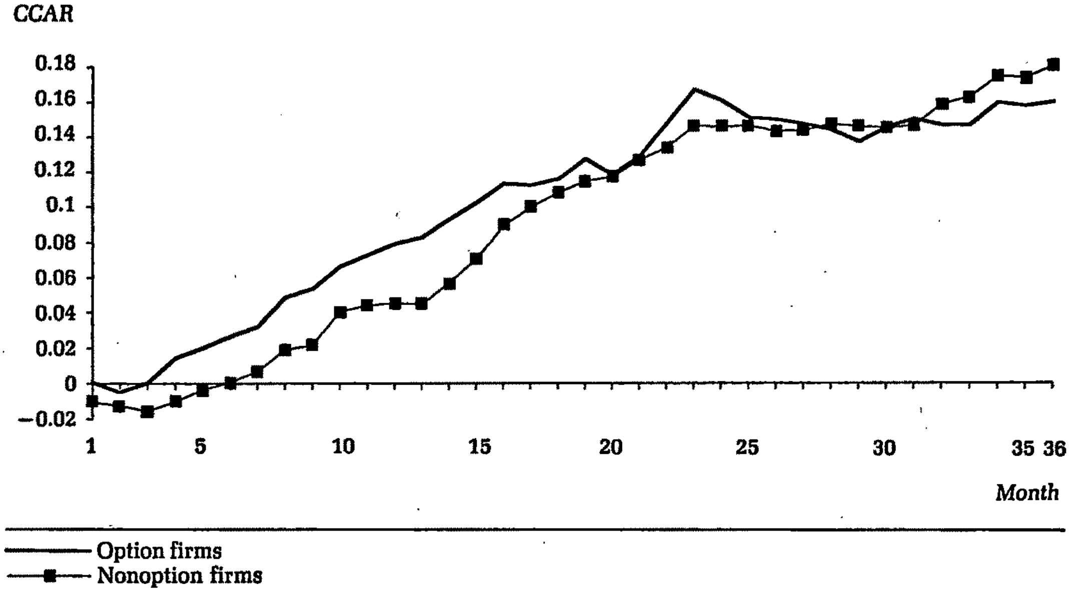
1. Firms must have December 31 fiscal year-ends to facilitate comparisons with previous studies that have imposed this restriction (e.g., Collins et al. 1987; Freeman 1987).
2. An option firm is deleted from the final option sample of year y if its 36-month return cumulation period includes any trading days prior to its listing date. This criterion ensures that the return cumulation period for each option firm in each sample year will not overlap the period in which the firm does not have exchange-listed options available in the market.

Results

Entire Sample. Figure 1 shows the drifts of CCAR for option versus nonoption firms when using the pooled data from 1980–1983. The results indicate that the 36-month CCAR of the nonoption firms is slightly greater than that of option firms (0.18 vs. 0.1599). This difference is not significant at the 0.1 level when using the empirical distribution generated from a randomization test.

Freeman (1987, table 6) suggests that small firms' cumulative returns are eventually larger than those of large firms in a 36-month accumulation period. The nonoption firms in this study are smaller than the option firms, but Freeman's "magnitude" hypothesis is not supported by the data. One potential explanation is that the small firms in Freeman's study tend to have higher earnings and price volatilities. In this

Figure 1
Comparison of Composite Cumulative Abnormal Returns (CCAR) Over 36 Months Preceding (1–12), During (13–24), and Following (25–36) Relevant Fiscal Years for Option and Nonoption Firms (1980–1983)



Note: Figure 1 reports the *risk-adjusted* composite cumulative abnormal returns (CCAR) for *the entire sample* over the period 1980–1983 (460 firm-year observations for the option sample and 375 for the nonoption sample). The CCAR scores are calculated for both option and nonoption portfolios that are constructed of long positions in good news firms and short positions in bad news firms. Good and bad news classifications are based on the direction of the change in income before extraordinary items and discontinued operations. The risk-adjusted abnormal returns are obtained from the market model and are calculated over the interval beginning 24 months prior to the end of the relevant fiscal year and ending 12 months after that fiscal year-end. The market model is estimated by OLS regressions over a 60-month period preceding the accumulation of abnormal returns. Firms must have at least 24 monthly returns available in the market model estimation period.

study, however, the option firms tend to have more volatile stock prices. In addition, as indicated in table 1, the earnings volatilities of the option and nonoption firms appear to be approximately the same. Thus, one would not expect the "magnitude" hypothesis to be supported.

The results, however, do indicate differences between option and nonoption firms in terms of the timeliness with which price changes anticipate earnings changes. As shown, the CCAR of the option firms begins its positive trend earlier than the CCAR of the nonoption firms (month 3 vs. month 6). Also, for option firms, the CCAR in month 12 (the end of the previous fiscal year) is 0.0793, which is about 50 percent of the total 36-month CCAR associated with the option firms' annual earnings changes. For nonoption firms, however, the CCAR in month 12 is 0.0454, which is only 25 percent of the total 36-month CCAR. The 25-point difference is significant at the 0.01 level when using the empirical distribution generated by a randomization test.[8] The results further indicate that 100 percent of the option firms' 36-month *CCAR* is realized by the end of the current fiscal year (month 24). At only 81 percent for the nonoption firms, the difference is significant at the 0.05 level with a randomization test. These results are consistent with hypothesis H2A and suggest that the security prices of option firms anticipate annual accounting earnings changes earlier than do those of nonoption firms.

The results also indicate that the option firms exhibit a different post-FYE price adjustment. The *CCAR* of the option firms declines from 0.1607 in month 24 to 0.1439 in month 28 (four months after the FYE), while the nonoption firms exhibit a slightly positive upward drift in *CCAR* from month 24 to month 28 (0.1458 to 0.1471). The *CCAR* of nonoption firms continues to drift up to 0.18 from month 28 to month 36. This evidence is consistent with hypothesis H2B and suggests a greater post-FYE drift for nonoption firms than for option firms.

Matched Subsamples and Control Portfolio. Consistent with the findings for the entire sample, the results for each matched subsample indicate that the proportion of the 36-month *CCAR* associated with the current year's earnings change that occurs in the previous fiscal year is significantly greater for the option firms than for the nonoption firms.[9] For the option firms, the proportion of the 36-month *CCAR* that is realized in month 12 (the end of the previous fiscal year) ranges from 53.97 percent (analyst subsample) to 60.11 percent (trading volume subsample). These proportions, however, range only from 17.93 percent (size subsample) to 28.95 percent (analyst subsample) for the nonoption firms. The differences are significant at the 0.05 level for all subsamples when using the empirical distributions generated from a randomization test. In the control portfolio, the proportion for option firms is 78.52 percent, but only 12.41 percent for nonoption firms. The difference is significant at the 0.01 level.

With respect to the post-FYE price adjustment, the results generally indicate a greater post-FYE drift for nonoption firms. For each subsample except size, the *CCAR* of the option firms either slightly declines or remains relatively constant from month 24 to month 36, but the *CCAR* of the nonoption firms exhibits a positive upward drift over the same period (on average, the magnitude of the upward drift is 4.17 percent). In the size subsample, both the option and nonoption firms experience an upward climb from month 24 to month 36. The magnitude of the upward drift, however, is still greater for the nonoption firms than for the option firms. In the control portfolio, the *CCAR* of the option firms declines from 0.1525 in month 24 to 0.1316 in month 36. The *CCAR* of the nonoption firms, however, exhibits a positive upward drift from month 24 to month 36 (0.1627 to 0.1970).

---

[8] The following randomization procedure is used to generate an empirical distribution of the difference between O_{12} and N_{12} with the pooled data over the period 1980–1983: (1) Option and nonoption firms in the good news portfolio are randomly shuffled to form two pseudo portfolios of good news, option and nonoption. Two pseudo portfolios of bad news are formed similarly. (2) Two pseudo samples are then formed by separating the option and nonoption firms out of the pseudo portfolios so that there are then good and bad news firms in both the option and nonoption pseudo samples. (3) Calculate O_{12} for the option pseudo sample and N_{12} for the nonoption pseudo sample. Let $D_{12}=O_{12}-N_{12}$. (4) Repeat the previous steps 1,000 times and use the distributions of the 1,000 values of D_{12} to assess the significance levels of the actual realization of D_{12}.

[9] The matching procedure for the long-window analysis is analogous to that for the short-window test. In each subsample, both Wilcoxon and two-sample *t*-tests fail to detect any significant differences in the matching variable between the matched option and nonoption firms.

IV. Concluding Remarks

This study documents differences in the price-earnings relation between firms with and without listed options. In particular, this study finds that the surprise associated with quarterly earnings announcements (as proxied by the abnormal stock return variability surrounding earnings announcements) is greater for nonoption firms than for option firms, and that the security prices of option firms anticipate annual accounting earnings changes earlier than those of nonoption firms.

Because option and nonoption firms differ on several firm-specific attributes and because these are likely to provide alternative explanations for the observed results, the empirical tests are repeated with a unilateral matching approach. The results for each matched subsample support the above findings, but are subject to a missing-variable problem. To mitigate this problem, a control portfolio that conservatively controls for the five proxy variables is constructed. With a relatively smaller sample size, the analysis again supports the above findings. In addition, multiple regression results indicate the incremental explanatory power of option trading over the other variables in explaining the differential content of earnings announcements for option and nonoption firms. However, given the low explanatory power of the variables in the regression, the inferences must be guardedly interpreted.

Finally, given (1) the lack of a formal theory concerning option trading and information flows and (2) the concern that the option exchanges choose firms with attributes that are likely to be correlated with measures of firms' information environments, it is difficult to conclude that option trading *causes* the differences in the price-earnings relation. This study considers the effects of five variables, but any other systematic differences between option and nonoption firms that are not included may provide alternative explanations for the observed results. Thus, the results simply suggest an *association* between option trading and returns behavior, both in conjunction with the release of earnings announcements and over periods during which the information contained in such announcements might be reaching the market.

References

Atiase, R. 1985. Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research* 23 (Spring): 21–36.

———. 1987. Market implications of predisclosure information: Size and exchange effects. *Journal of Accounting Research* 25 (Spring): 168–76.

Ball, R., and P. Brown. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6 (Autumn): 159–78.

Beaver, W. 1968. The information content of annual earnings announcements. *Journal of Accounting Research* 6 (Supplement): 67–92.

———, R. Lambert, and D. Morse. 1980. The information content of security prices. *Journal of Accounting & Economics* 2 (March): 3–28.

Black, F. 1975. Fact and fantasy in the use of options. *Financial Analysts Journal* 31 (July/August): 36–41 and 61–72.

Chiras, D., and S. Manaster. 1978. The information content of option prices and a test of market efficiency. *Journal of Financial Economics* 6 (June/September): 213–34.

Collins, D., S. Kothari, and J. Rayburn. 1987. Firm size and the information content of prices with respect to earnings. *Journal of Acounting & Economics* 9 (July): 111–38.

Cox, J. C., and M. Rubinstein. 1985. *Options Markets*. Englewood Cliffs, NJ: Prentice Hall.

Dempsey, S. 1989. Predisclosure information search incentives, analyst following, and earnings announcement price response. *The Accounting Review* 64 (October): 748–57.

Diamond, D., and R. Verrecchia. 1987. Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics* 18 (June): 277–311.

Dodd, P., N. Dopuch, R. Holthausen, and R. Leftwich. 1984. Qualified audit opinions and stock prices: Information content, announcement dates and concurrent disclosures. *Journal of Accounting & Economics* 6 (April): 3–38.

Freeman, R. 1987. The association between accounting earnings and security returns for large and small firms. *Journal of Accounting & Economics* 9 (July): 195–228.

Grant, E. 1980. Market implications of differential amounts of interim information. *Journal of Accounting Research* 18 (Spring): 255–69.

Jennings, R., and L. Starks. 1986. Earnings announcements, stock price adjustment, and the existence of option markets. *Journal of Finance* 41 (March): 107–25.

Lobo, G., and A. Mahmoud. 1989. Relationship between differential amounts of prior information and security return variability. *Journal of Accounting Research* 27 (Spring): 116–34.

Manaster, S., and R. Rendleman. 1982. Option prices as predictors of equilibrium stock prices. *Journal of Finance* 37 (September): 1043–57.

Marais, M. 1984. An application of the bootstrap method to the analysis of squared, standardized market model prediction errors. *Journal of Accounting Research* 22 (Supplement): 34–54.

Noreen, E. 1989. *Computer-Intensive Methods for Testing Hypotheses: An Introduction.* New York, NY: John Wiley & Sons.

O'Brien, P., and R. Bhushan. 1990. Analyst following and institutional ownership. *Journal of Accounting Research* 28 (Autumn): 55–76.

Patell, J. 1976. Corporate forecasts of earnings per share and stock price behavior: Empirical tests. *Journal of Accounting Research* 14 (Autumn): 246–76.

———, and M. Wolfson. 1979. Anticipated information releases reflected in call option prices. *Journal of Accounting & Economics* 1 (August): 117–40.

———, and ———. 1981. The *ex ante* and *ex post* price effects of quarterly earnings announcements reflected in option and stock prices. *Journal of Accounting Research* 19 (Autumn): 434–57.

Pincus, M. 1983. Information characteristics of earnings announcements and stock market behavior. *Journal of Accounting Research* 21 (Spring): 155–83.

Ross, S. 1976. Options and efficiency. *The Quarterly Journal of Economics* 90 (February): 75–89.

Shores, D. 1990. The association between interim information and security returns surrounding earnings announcements. *Journal of Accounting Research* 28 (Spring): 164–81.

Skinner, D. 1990. Options markets and the information content of accounting earnings releases. *Journal of Accounting & Economics* 13 (September): 191–211.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 385–399

The Explanatory Power of Earnings for Stock Returns

Norman Strong
Martin Walker
University of Manchester, England

SYNOPSIS AND INTRODUCTION: In a thorough review of market-based research on the information content of accounting earnings, Lev (1989) concludes that the explanatory value of earnings for stock returns, and therefore the usefulness of earnings disclosures, tends to be embarrassingly low. A number of nonmutually exclusive explanations have been advanced for these disappointing results, including: (1) poor specification of the estimating equation, such as a failure to allow for cross-sectional variation in the regression parameters; (2) inappropriate choice of the assumed proxy for expected earnings; (3) the availability of more timely sources of the value-relevant information in earnings statements (Beaver et al. 1980); and (4) poor informational properties (quality) of reported earnings because of biases induced by accounting measurement practices or creative "abuses" of the earnings measurement process.

Lev (1989) speculated that the last of these explanations was the most likely cause of the poor statistical performance consistently found in returns-earnings research. In contrast, the present study shows that a considerable improvement in statistical performance can be achieved by working with a more general specification of the returns-earnings relation. Lev's article has resulted in serious questioning of the contribution of market-based research, but we believe that the present study provides grounds for a more positive assessment.

We use a panel regression approach to examine the association between annual stock price returns and reported earnings figures of industrial companies in the United Kingdom. We combine several recent advances in

We are grateful to members of the Centre for Empirical Research in Accounting and Finance at the University of Manchester for comments on this research. An early version of the paper was presented at the 1991 Finance and Market Based Accounting Research Conference held at the University of Manchester; we are grateful for comments received from participants at that conference. John O'Hanlon and Peter Pope made a number of helpful suggestions on a previous version of the paper. We would particularly like to acknowledge the comments of the editor, an associate editor, and the anonymous referees; their comments resulted in a substantial improvement in the focus and exposition of the paper. The financial support of the ICAEW and the ESRC (grant number R000233813) is also gratefully acknowledged.

Submitted August 1991.
Accepted December 1992.

market-based acounting research design to produce a specification of the relation between earnings and price changes that subsumes the following key features:

1. Contemporaneous earnings yield is included in addition to the deflated first difference in earnings that is normally included in models of the returns-earnings relation.
2. Regression parameters are allowed to vary both cross-sectionally and over time.
3. Parameter values are allowed to vary across components of earnings to accommodate differences in the degree of persistence; in particular, we model the explanatory power resulting from attempts by accountants to distinguish extraordinary and exceptional items from the other components of earnings.

We introduce these features in a general model in a way that allows us to assess the incremental explanatory power of each individually as well as the joint effects of two or more combined. Each methodological improvement contributes significantly to our ability to explain security price changes, and we show that the best fit is achieved by incorporating all three features in a single general model. In moving from the standard model, which regresses a measure of abnormal returns on earnings changes, to the most general model, the adjusted *R*-squared increases from 0.10 to 0.38.

Key Words: *Stock returns, Earnings yield, Earnings disaggregation, Panel regression.*

Data Availability: *A list of the sample firms is available from the authors upon request. The rest of the data are available from sources identified in the text.*

THE article is organized as follows. We introduce the general earnings response model in section I and provide details of the sample of observations and definitions of variables in section II. Section III contains a statistical analysis of the variables included in the study. The principal results are reported in section IV, and we offer our conclusions in the final section.

I. Earnings Response Models

The standard model of the relation between abnormal returns and accounting earnings can be represented as:

$$CAR_{it}=a+b\frac{E_{it}-E_{it-1}}{V_{it-1}}+u_{it}, \tag{1}$$

where:

a = an intercept parameter to be estimated,
b = a slope parameter to be estimated (generally referred to in the literature as the earnings response parameter),

CAR_{it} = the cumulative abnormal return of company i in period t,
E_{it} = the earnings of firm i in period t,
V_{it-1} = the market value of equity at the start of period t, and
u_{it} = a conventional, mean zero, disturbance term.

Recent advances in the market-based accounting literature have attempted to extend this basic model in a number of ways. Theoretical work by Ohlson (1989) has suggested that earnings yield might be added as a second independent variable to allow for transitory components in earnings, and Easton and Harris (1991) have further highlighted the potential of the yield variable to explain stock returns.[1] In addition, a number of recent studies based on U.S. data have documented statistically significant variations in the earnings response coefficient, both across firms and over time. Kormendi and Lipe (1987) and Collins and Kothari (1989) show that earnings response coefficients are positively related to the persistence of earnings and exhibit cross-sectional and intertemporal variation. Easton and Zmijewski (1989), adopting a random coefficients regression model, and Board and Walker (1990), employing a fixed coefficients (dummy-variable) regression approach, support these findings. Other studies have shown that it may be possible to increase the explanatory power of earnings response models by identifying various components of earnings as separate independent variables (see, e.g., Lipe 1986).

Collectively, these considerations suggest the following generalization of equation (1):

$$CAR_{it} = a_t + \sum_{j=1}^{J} (b_i^j + b_t^j) \frac{E_{it}^j - E_{it-1}^j}{V_{it-1}} + \sum_{j=1}^{J} (c_i^j + c_t^j) \frac{E_{it}^j}{V_{it-1}} + u_{it}, \tag{2}$$

where:

a_t = a time-varying intercept parameter,
E_{it}^j = the jth component of the earnings of firm i in period t,[2]
b_i^j and b_t^j = firm-varying response parameters associated with the change in E_{it}^j,
c_i^j and c_t^j = firm-varying response parameters associated with the level of E_{it}^j,
c_i^j and c_t^j = are firm-varying response parameters associated with the level of E_{it}^j,

and other variables are as defined previously.

Equation (2) adds an intercept term that is allowed to vary over time to equation (1). In addition, the model disaggregates earnings into J components and allows the response parameters to vary across components and over time. Finally, equation (2)

---

[1] The standard earnings response model represented by equation (1) assumes that earnings follow a random walk. The potential value of including an earnings yield variable can be seen, if we instead assume that earnings follow a first-order autoregressive process of the form,

$$E_{it} - \mu = \phi(E_{it-1} - \mu) + e_{it}.$$

Then expected earnings in period t is, $v + \phi(E_{it-1} - v)$, and the returns-earnings relation becomes,

$$CAR_{it} = a + \frac{b(\phi - 1)\mu}{V_{it-1}} + b\phi \frac{E_{it} - E_{it-1}}{V_{it-1}} + b(1 - \phi) \frac{E_{it}}{V_{it-1}} + u_{it}.$$

[2] Throughout this article, superscripts are used exclusively to denote earnings components and the parameters associated with these components. Thus, E^2 denotes the second earnings component, not earnings squared.

incorporates a yield variable for each component of earnings, with the response parameter being allowed to vary across firms, over time, and across components.[3] In its most general form, the model requires estimation of $(J \times 7 + 3 \times T)$ parameters, where T is the number of years for which data are available. For 200 firms and 20 years of data per firm, the estimation of 1,460 parameters from 4,000 observations would be required.

We present empirical evidence for equation (2) and for several restricted versions that were carefully selected to highlight the incremental effects of: (1) allowing for cross-firm variation in the parameters, (2) introducing the yield variable into the specification, and (3) separately identifying the exceptional and extraordinary components of earnings. Following suggestions in a number of recent studies, we also present results for raw returns as well as for abnormal returns as the dependent variable.

II. Sample Selection and Data Definitions

The study is based on a random sample of 416 companies selected from the 1987 list of all U.K.-quoted industrial companies in Datastream.[4] From this starting sample, we selected all companies with December to April financial year-ends.[5] We carefully checked the sample to exclude any companies that changed their accounting year-end during the relevant period. Further companies were excluded by the following data requirements: (1) At least 10 consecutive years of accounting data had to be available on Datastream within the period 1971 through 1986 and (2) we also required a complete set of 16 years of returns data on the London Business School Share Price Database (LSPD). The 16 years correspond to the 10 years for which the accounting data was available, plus five years preceding the first year of accounting data (used to estimate the market model for the first year of the study) and one year following the last year of accounting data.

These criteria generated a final study sample of 146 U.K.-quoted companies and 2,036 observations. Though slightly biased toward the larger companies, the size distribution of the sample at the end of 1980, when all sample firms were trading, was reasonably representative of the population of firms recorded as being traded at the end of 1980 by the LSPD. Sixty percent of both the sample and the population of companies fell in the size range of £2 through £38 million ($4.8 through $91.2 million). However, the percentage of companies with a market capitalization of less than £2 millon was only 10 percent in the sample, compared to 20 percent in the entire population. At the other end of the size distribution, 30 percent of the sample firms had a market capitalization greater than £38 million, compared to only 20 percent in the population. The industrial composition of the sample is reported in table 1 and is reasonably representative of the population of industrial companies in the United Kingdom.

The values of two return metrics were computed for each company year of data.[6]

---

[3] An econometric restriction implicit in equation (2) is that the cross-firm and intertemporal sources of parametric variation are assumed to be additive in their effects.

[4] Datastream is an extensive on-line system of data bases covering, *inter alia*, domestic (United Kingdom) and international company accounts.

[5] In the regression analysis, the months January to April were grouped with the preceding December in defining period t.

[6] The statistical analysis used abnormal return calculated according to the capital asset pricing model as a third return metric. The results were virtually identical to those for the market model abnormal return and are not separately reported.

Table 1
Industrial Composition of the Study Sample

Industrial Classification	N	*Percent*
Industrial and building material	17	11.6
Contracting, construction	9	6.2
Electricals and electronics	4	2.7
Miscellaneous mechanical engineering	16	11.0
Other capital goods	21	14.4
Brewering, catering, and leisure	9	6.2
Food	7	4.8
Newspaper, publishing, and printing	10	6.8
Stores and other retailing	10	6.8
Health and household products	7	4.8
Chemicals and oil	5	3.4
Shipping and transport	10	6.8
Miscellaneous industrials	7	4.8
Overseas trade	6	4.1
	2	1.4
Total	146	100

MMCAR is the cumulative monthly market model abnormal return from May of year t through April of year $t+1$, and *RETURN* is the cumulative monthly stock return from May of year t through April of year $t+1$.[7] Both return metrics implicitly assume that the earnings news for year t (for a December year-end company) is released no earlier than May of year t and no later than April of year $t+1$.

The empirical analysis focuses on three components of earnings: earnings before exceptional and extraordinary items, denoted E^1 and referred to as "pre-exceptional earnings"; exceptional earnings, denoted E^2; and extraordinary earnings, denoted E^3. The following important aggregations of these earnings components are also examined:

$$\text{Ordinary earnings} \equiv E^1 + E^2, \qquad (3)$$

and

$$\text{All-inclusive earnings} \equiv E^1 + E^2 + E^3. \qquad (4)$$

In the United Kingdom, ordinary earnings comprise the reported earnings number on which earnings per share calculations are currently based.[8]

---

[7] The dates for estimating benchmark returns, for cumulating (abnormal) returns, and for deflating earnings components refer to December year-end companies. For company financial years ending January to April, dates are successively adjusted forward by one month. The market model abnormal return was calculated by fitting the market model to the continuously compounded company and market returns (*Financial Times* All Share Index) for the 60 months immediately prior to year t. The abnormal returns from May of year t to April of year $t+1$ were then derived as the prediction errors from the market model, conditional on the realized value of the market index for that month. All return data were taken from the LSPD.

[8] The distinctions between pre-exceptional, exceptional, and extraordinary earnings were established in the early 1970s by the United Kingdom's Accounting Standards Committee (ASC) and were formalized in Statement of Standard Accounting Practice 6 (SSAP6) in 1974. Extraordinary items were defined as "those items which derive from events or transactions outside the ordinary activities of the business and which are both material and expected not to recur frequently or regularly. They do not include those items which, though

Corresponding to the definitions, the following five accounting variables were assembled from Datastream company accounts data for each company year:

E^1+E^2 = Ordinary earnings (Datastream item 175),
E^2 = Exceptional items (Datastream item 194),
E^3 = Extraordinary items (Datastream item 193),
E^1 = Earnings before Exceptional Items (Datastream items 175 minus 194), and
$E^1+E^2+E^3$ = All-inclusive earnings (Datastream items 175 plus 193).

For each of these earnings components, we then defined and calculated the "deflated first difference" as the value of the component in year t minus the value in year $t-1$ divided by the market capitalization on April 30 of year t (May 31 for January year-end companies, etc.). We also calculated an earnings yield variable for each component as the value of the component reported for year t divided by the market capitalization on April 30 of year t. Dummy variables were constructed to allow the regression parameters to vary both across firms and over time in accordance with equation (2).[9] The time dummies allow for a different response parameter for each year (corresponding to an average coefficient for all firms in that year). The (time-invariant) firm dummies allow for a different response parameter for each firm.[10]

III. Data Analysis

Before conducting the panel regressions, we completed a number of separate analyses on the materiality of the earnings components, the correlation structure of the main variables involved in the study, and the time-series properties of the three earnings components.

Materiality of the Earnings Components in Relation to Market Capitalization

This was assessed by examining the key descriptive statistics of the three components of all-inclusive earnings yield (see table 2). The pre-exceptional earnings yield averaged about 13 percent over the sample period, with a standard deviation of 20 percent. There was no marked skewness. The median values and the 25th and 75th per-

---

exceptional on account of size or incidence (and which therefore may require separate disclosure), derive from the ordinary activities of the business. Neither do they include prior year items simply because they relate to a prior year." This definition implicitly considers exceptional items as arising from the ordinary activities of the business and requiring separate disclosure because of their size or incidence.

Successive surveys of published accounts showed that SSAP6 succeeded in rapidly establishing the required practice of separate disclosure of extraordinary and exceptional items and the practice of calculating earnings per share figures according to earnings after exceptional but before extraordinary items (see, e.g., ICAEW 1977, sec. 7). However, several surveys also showed considerable inconsistencies in the way similar items were accounted for by different companies, the main difficulty being that of deciding whether a particular item derived from the ordinary activities of the business. In an attempt to reduce the incidence of such inconsistencies, the ASC issued a revised version of SSAP6 in 1986. To avoid complications caused by the changed definition of extraordinary items, we restrict our attention to the period preceding the revision.

[9] Judge et al. (1985, chap. 13) gives a helpful discussion of the basic technique.

[10] We found that 37 companies reported extraordinary items in fewer than six years. In regressions with extraordinary items separately identified, the corresponding firm-varying regression parameters were restricted to be constant across these companies. A similar procedure was followed for 21 companies that reported exceptional items in fewer than six years.

Table 2
Descriptive Statistics for the Three Earnings Yield Variables

Components of All-Inclusive Earnings	*Maximum*	*75th Percentile*	*Median*	*25th Percentile*	*Minimum*	*Mean*	*Standard*	*Skewness*
Pre-exceptional earnings yield	2.37	0.17	0.11	0.07	−1.46	0.13	0.20	1.15
Exceptional earnings yield	0.85	0.01	0.00	0.00	−0.51	0.01	0.05	5.55
Extraordinary earnings yield	1.10	0.00	0.00	−0.01	−2.46	−0.02	0.14	−7.01

Note: All the components are deflated by the opening equity capitalization; N=2,036.

Table 3
Median Time-Series Correlations of *MMCAR* and *RETURN* with the Independent Variables

Independent Variables	*Dependent Variable* MMCAR	 *RETURN*
Pre-exceptional earnings yield	0.39	0.34
Pre-exceptional first difference	0.44	0.37
Exceptional earnings yield	0.00	−0.01
Exceptional first difference	−0.04	−0.01
Extraordinary earnings yield	0.01	0.06
Extraordinary first difference	0.17	0.16

Note: The variables comprise the three components of all-inclusive earnings, each expressed in levels and first-difference form. Each is deflated by opening equity capitalization. N=all 146 companies.

centiles of the exceptional and extraordinary earnings yields show that many of these reported values are close to zero—50 percent are less than or equal to 1 percent in absolute value. It is also noteworthy that the extraordinary items have a negative skew, in contrast to the positive skew of exceptional items, which suggests a tendency to report large positive items as exceptional and large negative items as extraordinary.

Correlation Structure

The median time-series correlations of *MMCAR* and *RETURN* with the three earnings yield and three first-difference variables are given in table 3. The median value of the correlations for the two pre-exceptional variables are both appreciably greater than zero. Moreover, the 25th percentiles of the distributions are also positive. In contrast, the correlations of the variables for the extraordinary and exceptional items tend to be much lower. The median correlations for the two exceptional earnings variables are close to zero; although positive for the two extraordinary earnings variables, they are appreciably lower than the corresponding correlations for pre-exceptional earnings.

The median time-series correlation matrix for the six independent variables (table 4) exhibits four significant features:

1. The median correlations between the exceptional and extraordinary variables (levels and first differences) are all close to zero.
2. The median correlation between the first difference of pre-exceptional earnings and the pre-exceptional earnings yield, at 0.64, is moderately high.
3. The two exceptional variables have a tendency to be negatively correlated with the two pre-exceptional earnings variables. The extraordinary variables, however, exhibit a positive association. These are the kind of associations one would expect if exceptional and extraordinary items are used to "smooth" reported earnings per share.
4. The median correlation between the yield and difference variables is high for both exceptional and extraordinary items. Multicollinearity could prove to be a problem here.[11]

---

[11] In particular, it would be quite possible for the yield and differences variables to be jointly significant but for both individual *t*-statistics to be insignificant.

Table 4
Median Time-Series Correlation Matrix of the Main Independent Variables

	Pre-Exceptional Difference	*Exceptional Yield*	*Exceptional Difference*	*Extraordinary Yield*	*Extraordinary Difference*
Pre-exceptional yield	0.64	−0.36	−0.22	0.14	0.09
Pre-exceptional difference		−0.39	−0.52	0.04	0.11
Exceptional yield			0.74	0.00	0.01
Exceptional difference				0.01	0.01
Extraordinary yield					0.74

Note: The variables comprise the three components of all-inclusive earnings, each expressed in levels and first-difference form. Each is deflated by opening equity capitalization. N=all 146 companies.

Time-Series Properties

The first-order serial correlations of the three main earnings variables (undeflated levels) are shown in table 5. In contrast to pre-exceptional earnings, which exhibit strong positive serial correlation with a median value of 0.63, the exceptional and extraordinary items exhibit a median serial correlation close to zero. Moreover, the 75th percentiles of the distributions for the exceptional and extraordinary items lie below the 25th percentile of the corresponding distribution for pre-exceptional earnings. As an alternative test for persistence, we estimated the first-order serial correlations for the corresponding first differences. The median first-difference correlation for pre-exceptional earnings is −0.01, close to zero. For exceptional and extraordinary earnings, the corresponding median correlations are −0.41 and −0.35, respectively, which shows strong evidence of mean reversion.[12]

The results suggest that pre-exceptional earnings exhibit a high degree of persistence but that the other two components are highly transitory. In conformance with the ideas of Kormendi and Lipe (1987) and Collins and Kothari (1989), these persistence comparisons lead us to hypothesize that, on average, the relation between stock price movements and extraordinary items should be broadly similar to the corresponding relation for exceptional items. Moreover, both of these relations should differ from the corresponding relation for pre-exceptional earnings. Because of these considerations, our general model separately identifies the exceptional and extraordinary components of earnings by allowing the response parameters to vary across components.

IV. Regression Results

The main findings of the study are reported in two subsections. The first shows the improvements in explanatory value achieved by successively moving toward the general econometric specification previously outlined. The second examines the signs and significance levels of the parameter estimates generated by our most general econometric model.

[12] We also employed nonparametric procedures to test whether the distribution of the first-order serial correlations differed between the three earnings variables. Kolmogorov-Smirnov and Kuiper statistics, based on the empirical distribution function, rejected the null hypothesis that the distribution of first-order serial correlations for pre-exceptional earnings is the same as that for either exceptional or extraordinary earnings, at a significance level of 0.1 percent. The same test statistics were unable to reject the null hypothesis that the distributions of first-order serial correlations for exceptional and extraordinary items are the same (at well above standard significance levels).

Table 5
First-Order Serial Correlations of the (Undeflated) Earnings Variables

	25th Percentile	*Median*	*75th Percentile*
Pre-exceptional earnings	0.31	0.63	0.85
Exceptional earnings	−0.17	0.01	0.24
Extraordinary earnings	−0.09	0.05	0.29

Table 6
A Summary of the Differences Between Models for Which Adjusted R^2 are Reported in Table 7

Model Number	*Time-Varying Parameters*	*Firm-Varying Parameters*	*Levels Variable Included*	*Earnings Disaggregated*
M0	No	No	No	No
M1	Yes	No	No	No
M2	Yes	No	Yes	No
M3	Yes	No	No	Yes—$\{E^1\}, \{E^2\}, \{E^3\}$
M4	Yes	Yes	No	No
M5	Yes	No	Yes	Yes—$\{E^1\}, \{E^2\}, \{E^3\}$
M6	Yes	Yes	No	Yes—$\{E^1\}, \{E^2\}, \{E^3\}$
M7	Yes	Yes	Yes	No
M8	Yes	Yes	Yes	Yes—$\{E^1\}, \{E^2+E^3\}$
M9	Yes	Yes	Yes	Yes—$\{E^1+E^2\}, \{E^3\}$
M10	Yes	Yes	Yes	Yes—$\{E^1\}, \{E^2\}, \{E^3\}$

Model Selection

The 11 models for which results are reported are presented in figure 1 and summarized in table 6. The starting point for our analysis is M0, the basic earnings response model given by equation (1). This model was fitted to the entire (pooled) sample of observations with the first difference in all-inclusive earnings as the independent variable. The resulting $\bar{R}$-squared values are reported in table 7 and provide a benchmark for comparison with the more refined models, M1 to M10.

The first refinement is given by model M1, which allows the regression parameters of the basic model to vary over time. A number of researchers have established theoretical grounds for time-series variation in the earnings response parameter (see, e.g., Board and Walker 1990; Collins and Kothari 1989). In addition, the earnings response parameter will vary over time if the expectation benchmark for earnings is more accurate in some years than in others.

A comparison of the findings for M1 with those for M0 reveals a marked improvement in $\bar{R}$-squared under both *MMCAR* and *RETURN*. The improvement for *MMCAR*

Figure 1
A Summary of the Panel Regression Models

Given below is the full set of regression equations used in the research. All equations were estimated with both annual cumulative abnormal return (*MMCAR*) and annual unadjusted return (*RETURN*) as the dependent variable (denoted below as R_{it}). The components of earnings identified are pre-exceptional earnings, E^1, exceptional earnings, E^2, and extraordinary earnings, E^3. Where appropriate, the equations are given in increasing order of generality.

$$R_{it}=a+b\frac{\sum_{j=1}^{3}(E^j_{it}-E^j_{it-1})}{V_{it-1}}+u_{it} \qquad \text{(M0)}$$

$$R_{it}=a_t+b_t\frac{\sum_{j=1}^{3}(E^j_{it}-E^j_{it-1})}{V_{it-1}}+u_{it} \qquad \text{(M1)}$$

$$R_{it}=a_t+b_t\frac{\sum_{j=1}^{3}(E^j_{it}-E^j_{it-1})}{V_{it-1}}+c_t\frac{\sum_{j=1}^{3}E^j_{it}}{V_{it-1}}+u_{it} \qquad \text{(M2)}$$

$$R_{it}=a_t+\sum_{j=1}^{3}b^j_t\frac{E^j_{it}-E^j_{it-1}}{V_{it-1}}+u_{it} \qquad \text{(M3)}$$

$$R_{it}=a_t+(b_t+b_i)\frac{\sum_{j=1}^{3}(E^j_{it}-E^j_{it-1})}{V_{it-1}}+u_{it} \qquad \text{(M4)}$$

$$R_{it}=a_t+\sum_{j=1}^{3}b^j_t\frac{E^j_{it}-E^j_{it-1}}{V_{it-1}}+\sum_{j=1}^{3}c^j_t\frac{E^j_{it}}{V_{it-1}}+u_{it} \qquad \text{(M5)}$$

$$R_{it}=a_t+\sum_{j=1}^{3}(b^j_t+b^j_i)\frac{E^j_{it}-E^j_{it-1}}{V_{it-1}}+u_{it} \qquad \text{(M6)}$$

$$R_{it}=a_t+(b_t+b_i)\frac{\sum_{j=1}^{3}(E^j_{it}-E^j_{it-1})}{V_{it-1}}+(c_t+c_i)\frac{\sum_{j=1}^{3}E^j_{it}}{V_{it-1}}+u_{it} \qquad \text{(M7)}$$

$$R_{it}=a_t+(b^1_t+b^1_i)\frac{E^1_{it}-E^1_{it-1}}{V_{it-1}}+(b_t+b_i)\frac{\sum_{j=2}^{3}(E^j_{it}-E^j_{it-1})}{V_{it-1}}$$
$$+(c^1_t+c^1_i)\frac{E^1_{it}}{V_{it-1}}+(c_t+c_i)\frac{\sum_{j=2}^{3}E^j_{it}}{V_{it-1}}+u_{it} \qquad \text{(M8)}$$

$$R_{it}=a_t+(b_t+b_i)\frac{\sum_{j=1}^{2}(E^j_{it}-E^j_{it-1})}{V_{it-1}}+(b^3_t+b^3_i)\frac{E^3_{it}-E^3_{it-1}}{V_{it-1}}$$
$$+(c_t+c_i)\frac{\sum_{j=1}^{2}E^j_{it}}{V_{it-1}}+(c^3_t+c^3_i)\frac{E^3_{it}}{V_{it-1}}+u_{it} \qquad \text{(M9)}$$

$$R_{it}=a_t+\sum_{j=1}^{3}(b^j_t+b^j_i)\frac{E^j_{it}-E^j_{it-1}}{V_{it-1}}+\sum_{j=1}^{3}(c^j_t+c^j_i)\frac{E^j_{it}}{V_{it-1}}+u_{it} \qquad \text{(M10)}$$

Table 7
Adjusted R-Squared Values of the Models

	Dependent Variables	
Regression Models	MMCAR	*RETURN*
M0	0.1028	0.0659
M1	0.1943	0.3753
M2	0.2083	0.3920
M3	0.2195	0.3940
M4	0.2386	0.4141
M5	0.2412	0.4134
M6	0.3059	0.4758
M7	0.2737	0.4796
M8	0.3142	0.5069
M9	0.3359	0.5240
M10	0.3823	0.5454

Note: The independent variables comprise various disaggregations of all-inclusive earnings, in first difference or levels form or both. These are deflated by opening equity capitalization, and various degrees of temporal and cross-sectional variation are modeled.

is broadly consistent with the findings of previous studies based on U.S. data. A more dramatic increase in $\bar{R}$-squared occurs with *RETURN* as the dependent variable because, in addition to the factors that cause the regression parameters to vary under *MMCAR*, a considerable proportion of the time-series variation in individual company returns is "explained" by the market; in model M1, the temporal variation in the market return is captured by the time-varying intercept.

The remaining results for models M2 to M10 in table 7 incorporate various combinations of refinements over and above those captured by model M1. These results allow us to assess the individual and joint effects of incorporating (1) firm variation in the independent variable response parameters, (2) earnings yield variables, and (3) earnings disaggregation. We first consider the individual effects of these three additional refinements, then the joint effects of incorporating two or more of the refinements.

Comparing the findings for models M2, M3, and M4 with those for model M1 shows that moderate improvements in $\bar{R}$-squared are achieved by all three methodological refinements under both *MMCAR* and *RETURN*. Conventional analysis of covariance (*F*-) tests confirmed that the increases in $\bar{R}$-squared of models M2, M3, and M4 over M1 are significant at the 1 percent level under both *MMCAR* and *RETURN*.[13]

Models M5, M6, and M7 show the joint effect of introducing two of the three model refinements. Analysis of covariance tests were performed to compare these models with their nested counterparts in models M2, M3, and M4.[14] These tests showed that significant improvements in explanatory power at the 1 percent level are achieved for both dependent variables in M5 over M2 and M3, M6 over M3 and M4, and M7 over M2 and M4.

---

[13] See Harvey (1990, 61–64) for further details.

[14] A model is a nested version of another if it can be derived by imposing appropriate restrictions on the parameters of the other model. See Judge et al. (1985, chap. 21), for details.

The effect of incorporating all three methodological refinements into our general model was assessed by comparing model M10 with models M5, M6, and M7. Analysis of covariance tests rejected M5, M6, and M7 in favor of model M10 at a significance level of 1 percent for both *MMCAR* and *RETURN*. Thus, adding any one of the three methodological refinements yields a further significant improvement in explanatory power. Overall, these tests identify M10 as the most appropriate model of the relation between returns and earnings. Moreover, the improvement in explanatory power generated by moving from the basic earnings response model, M0, to the general model, M10, suggests that the low $\bar{R}$-squared values reported in the research surveyed by Lev (1989) can, to a significant extent, be attributed to the restrictive research designs in early studies.

Finally, for sake of completeness, we investigated the comparative explanatory value of three alternative forms of earnings disaggregation via comparisons of model M10 with models M8, which aggregates exceptional and extraordinary items but separates these from pre-exceptional earnings, and M9, which distinguishes between ordinary and extraordinary earnings. Examination of M8 is motivated by our earlier results on the time-series properties of the three earnings components; examination of M9 follows from the standard accounting practice in the United Kingdom of separating extraordinary items from ordinary earnings. Conventional analysis of covariance tests selected M10 over both M8 and M9 under both *MMCAR* and *RETURN* at a significance level of 1 percent. These results suggest that the market-pricing process for exceptional items is different from the pricing process for both pre-exceptional earnings and extraordinary items.[15]

Parameter t-*Values and Hypothesis Tests*

Since the number of parameter values generated by M10 is large (869), the analysis focuses on hypothesis tests relating to the average value of the parameters across firms and over time. Table 8 shows the significance of the average values, across firms and time, of the regresssion slope parameter estimates generated by fitting model M10.[16] The results are similar under both dependent variables.[17] The significance of both pre-exceptional earnings yield and pre-exceptional first difference suggests that the specification of the conventional earnings response model is improved by incorporating a yield variable.

None of the other four variables is significant at conventional levels under either *MMCAR* or *RETURN*, a result apparently contradicting the earlier finding that model M10 provides a better fit to the data than model M8. One possible reason for this lack of significance is the multicollinearity noted in the discussion of table 4. To test this, we conducted further *t*-tests for the *joint* significance of the yield and first-difference pa-

---

[15] For completeness, we also ran the alternative regression to M8 and M9 by aggregating pre-exceptional earnings (E^1) and extraordinary items (E^3) but separating exceptional earnings (E^2). The $\bar{R}$-squared for this regression was 0.3066 for *MMCAR* and 0.5012 for *RETURN*.

[16] Bernard (1987, 13) discusses the severe problems for hypothesis testing that arise from cross-sectional dependency in the residuals when models impose cross-sectional homogeneity in the response coefficient. By allowing for cross-sectional variation, M10 should substantially reduce the problem of cross-sectional dependence.

[17] The remaining problem of serial correlation in the residuals could lead to upwardly biased *t*-values when *RETURN* is the dependent variable. However, the similarity of the findings for *RETURN* and *MMCAR* is reassuring.

Table 8
Average Values of the Regression Slope Parameters for Model M10

Independent Variable(s)	*Dependent Variable* MMCAR	 *RETURN*
Pre-exceptional earnings yield	1.59**	1.67**
Pre-exceptional earnings difference	1.85**	1.92**
Exceptional earnings yield	1.48	0.49
Exceptional earnings difference	2.51	3.85
Extraordinary earnings yield	−0.37	0.90
Extraordinary earnings difference	0.60	−0.60
Exceptional earnings yield and difference	2.00*	2.17**
Extraordinary earnings yield and difference	0.11	−0.05

Note: The test determines whether the average firm slope parameter corresponding to a particular independent variable (or variables) is significantly different from zero when the slope parameter for an individual firm is given by its firm-varying component for those years the firm was in the sample. All tests used the full variance-covariance matrix of parameter estimates from model M10.

rameters for both exceptional and extraordinary earnings (see table 8). The important additional result to emerge is rejection of the null hypothesis that the exceptional yield and difference variables taken jointly are, on average, insignificantly different from zero.[18]

V. Conclusions

The evidence in this study suggests that a significant improvement in the statistical performance of models of earnings and returns can be achieved by allowing for time-series and cross-sectional variation in the regression parameters, by including an earnings yield variable along the lines suggested by Ohlson (1989), and by partitioning all-inclusive earnings into pre-exceptional, exceptional, and extraordinary components. The evidence that the pre-exceptional earnings yield variable is statistically significant on average further suggests that pre-exceptional earnings exhibit both permanent and transitory features and that models of the relation between earnings and returns that focus exclusively on the deflated first difference of earnings are misspecified.

---

[18] Shortly before publication we found that Datastream Item 175 excluded certain exceptional items in some of our sample company years. We re-estimated the models using Datastream Item 182, which includes all exceptional items. Model M10 still emerged as the best model with $\bar{R}$-squared values marginally higher than those reported in table 6. The pre-exceptional earnings yield variable was more significant than reported above and was the only variable to be statistically significant on average. Further details are available from the authors on request.

References

Beaver, W. H., R. Lambert, and D. Morse. 1980. The information content of security prices. *Journal of Accounting & Economics* 2 (March): 3–28.

Bernard, V. L. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* 25 (Spring): 1–48.

Board, J. L. G., and M. Walker. 1990. Intertemporal and cross-sectional variation in the association between unexpected rates and return and abnormal returns. *Journal of Accounting Research* 28 (Spring): 182–92.

Christie, A. A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting & Economics* 9 (December): 231–58.

Collins, D., and S. Kothari. 1989. An analysis of intemporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting & Economics* 11 (2/3): 143–81.

Easton, P., and M. E. Zmijewski. 1989. Cross-sectional variation in the stock market response to accounting earnings announcements. *Journal of Accounting & Economics* 11 (2/3): 117–41.

——, and T. S. Harris. 1991. Earnings as an explanatory variable for returns. *Journal of Accounting Research* 29 (Spring): 19–36.

Harvey, A. 1990. *The Econometric Analysis of Time Series.* 2d ed. New York: Philip Allen.

Institute of Chartered Accountants in England and Wales. 1977. *Survey of Published Accounts 1977.* London: Institute of Chartered Accountants in England and Wales. ICAEW.

Judge, G. G., W. E. Griffiths, R. Carter Hill, R. H. Lütkepohl, and Tsoung-Chao Lee. 1985. *The Theory and Practice of Econometrics.* 2d ed. New York: John Wiley and Sons.

Kormendi, R., and R. Lipe. 1987. Earnings innovations, earnings, persistence, and stock returns. *Journal of Business* 60 (3): 323–45.

Lev, B. 1989. On the usefulness of earnings and earnings research: Lessons and directions from two decades of empirical research. *Journal of Accounting Research* 27 (Supplement): 153–201.

Lipe, R. C. 1986. The information contained in the components of earnings. *Journal of Accounting Research* 24 (Supplement): 37–64.

Ohlson, J. A. 1989. Accounting earnings, book value, and dividends: The theory of the clean surplus equation (part 1). Working paper, Columbia University, New York.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 400–410

The Effect of Budget Emphasis and Information Asymmetry on the Relation Between Budgetary Participation and Slack

Alan S. Dunk
University of Western Sydney, Nepean

SYNOPSIS AND INTRODUCTION: A major concern in the literature is that participation by subordinates may result in the generation of slack budgets (Antle and Eppen 1985). In one of the earliest studies, Williamson (1964) concluded that subordinate managers will try to influence the budget-setting process and obtain slack budgets. In conformance with Merchant (1985a), Lukka (1988), and Young (1985), budgetary slack is defined as the express incorporation of budget amounts that make it easier to attain. Managers may build slack into budgets by strategies that understate revenues and overstate costs (Schiff and Lewin 1970).

Whether budgetary slack is a likely outcome in all participatively set budgets is a matter of conjecture. Lukka (1988) argued that a high degree of participation gives subordinate managers the opportunity to contribute directly to the creation of slack, and vice versa. However, the link between participation and slack is equivocal, since Cammann (1976), Merchant (1985a), and Onsi (1973) provide evidence that participation may lead to a reduction in slack, which can be attributed to the positive communication between managers so that subordinates feel less pressure to create slack.

The literature proposes a link between participation and budgetary slack through two variables: superiors' budget emphasis in their evaluation of subordinate performance, and the degree of information asymmetry between superiors and subordinates. When participation, budget emphasis, and information asymmetry are high (low), slack will be high (low).

Thanks are due to George Cooney, Graeme Harrison, Ken Merchant, Ken Moores, David Otley, and to participants at workshops at the Flinders University of South Australia and the University of Western Sydney, Nepean, for their comments on previous drafts of this project. The helpful comments of the anonymous referees are also gratefully acknowledged. Special thanks are reserved for Peter Brownell, whose contribution and support was invaluable.

Submitted March 1991.
Accepted December 1992.

For this study, samples of managers were drawn from manufacturing organizations in the Sydney, Australia, metropolitan area. Measures of budgetary slack and information asymmetry were developed. Support was found for low (high) slack when the predictors are high (low).

Key Words: *Budgetary slack, Information asymmetry, Budgetary participation, Budget emphasis.*

Data Availability: *The data are available from the author.*

THE remainder of the article is organized as follows. In the following section, the literature is reviewed and the theory is developed. The next section details the method, and is followed by a presentation of the results. Finally, the conclusions, implications, and limitations of the study are discussed.

I. Previous Literature and Theoretical Development

The Relation Between Participation and Budgetary Slack

Magee (1980) proposed that the expected payoff to principals (superiors) could be enhanced with access to local information held by agents (subordinates) prior to setting the budget. Local information is an example of information asymmetry (Baiman and Evans 1983; Coughlan and Schmidt 1985; Penno 1984), which arises when the subordinate has information relevant to the decision process associated with budgeting (Kren and Liao 1988). Information asymmetry exists only when subordinates' information exceeds that of their superiors.

Subordinates' participation in the budget-setting process gives principals the opportunity to gain access to local information (Baiman 1982; Baiman and Evans 1983; Magee 1980) which allows subordinates to communicate or reveal some of their private information that may be incorporated into the standards or budgets against which their performance would be assessed (Baiman and Evans 1983). But agents may misrepresent or withhold some of their private information, which could lead to budgets with slack (Christensen 1982; Merchant 1985b; Pope 1984; Young 1985). Hence, while participation of subordinates may give superiors access to private information, information asymmetry may enable slack to be increased in participatively set budgets.

Budget Emphasis. The primary argument for agents' efforts to build slack in their budgets is to enhance their compensation prospects. If subordinates perceive their rewards as dependent on budget attainment, they may try to build slack into their budgets through the participation process (see, e.g., Lowe and Shaw 1968; Schiff and Lewin 1968, 1970; Waller 1988). Thus, budget emphasis in performance evaluation induces the creation of budgetary slack (Baiman and Lewis 1989).

The Interaction of Information Asymmetry, Budget Emphasis, and Participation. If budget emphasis and information asymmetry are high, subordinates are likely to attempt to negotiate slack budgets (see, e.g., Baiman and Lewis 1989; Christensen 1982). If information asymmetry is low and budget emphasis is high, subordinates will have the incentive to create slack, but will not be in a position to obtain it (Chow et al. 1988; Penno 1984; Waller 1988). In all cases, participation in the budget-setting process will need to be high for the subordinate to have the opportunity to build in slack (Lukka

1988). When participation is low, the prospects for subordinates' building slack into their budgets are restricted, irrespective of the level of budget emphasis and information asymmetry. Thus, the literature suggests that high (low) participation with high (low) information asymmetry and high (low) budget emphasis is likely to give rise to high (low) budgetary slack. The hypothesis in its null form is stated as follows:

> H_0: There is no interaction between budgetary participation, information asymmetry, and budget emphasis that affects budgetary slack.

II. Method

To test this hypothesis, measures of budget emphasis and budgetary participation were drawn from the literature, but measures of budgetary slack and information asymmetry required development.

Sample Selection

Companies selected for inclusion in the study were drawn randomly from the population of manufacturing organizations with more than 100 employees located in the Sydney, Australia, metropolitan area using the *Kompass Australia*, (1988) register of firms. It was expected that companies with fewer employees were unlikely to have clearly defined areas of responsibility for managers. A total of 61 firms were selected for inclusion in the sample. No restriction was placed on the areas of responsibility from which managers could be drawn for sampling purposes.[1]

Each firm was contacted by telephone, and 118 managers were personally requested to contribute to the study. A questionnaire with a cover letter and reply-paid self-addressed envelope was mailed to each manager, and a telephone follow-up was made to promote a higher response rate. This follow-up also verified that the targeted managers had responded to the questionnaire themselves. Responses were received from 79 managers, an overall response rate of 67 percent. The respondents' mean age was 42 years, their average length of experience was 13 years, they had held their current positions on average for 4.5 years, and the mean number of employees in their areas of responsibility was 102.

Variable Measurement

An Analysis of the Draft Measure of Budgetary Slack. The draft six-item measure of budgetary slack is reported in appendix A. The items in the instrument are based on the definition of budgetary slack employed in this study. As such, they focus on the ease with which budgetary targets can be achieved. The items are anchored by (1) strongly disagree and (7) strongly agree.[2]

A factor analysis of the draft measure indicated that items 1 and 3 did not load on the factor at the 0.5 level, which was the criterion for item inclusion; therefore, both items were subject to exclusion from the instrument. Since item 3 had a much lower loading than item 1, the decision was made to remove only item 3 first, and to then evaluate the findings of a further factor analysis based on the remaining five items. In this analysis, all five items loaded on a single factor. Because item 1 again did not load on the factor at the 0.5 level, it was excluded from the instrument. A factor analysis of the remaining four items (2, 4, 5, and 6) indicated that they loaded above 0.5 on a single factor, which had an eigenvalue of 1.3878 and explained 34.7 percent of the variation in the underlying variable. In addition, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.5827.[3]

---

[1] The findings of experimental research suggest that subjects' risk preferences affect the creation of slack (see, e.g., Chow et al. 1988; Waller 1988; Young 1985). While Waller (1988) controlled risk preferences experimentally, this study relies upon random sampling to control for the possible confounding effects of risk preferences on the results (Dominowski 1980; Monette et al. 1986).

[2] The study relies upon managers' perceptions of the level of slack in their budgets. Although more objective data on slack levels might be desirable, the means for generating this type of data are difficult to ascertain.

Based on these findings, a measure of budgetary slack was construced from items 2, 4, 5, and 6. Although the correlation between this four-item measure and the originally proposed six-item measure is 0.932 ($\underline{p} < 0.001$), and the Cronbach alpha is virtually identical to that of the originally proposed measure (0.6813), the findings enabled a more sharply focused measure.[4] Descriptive statistics for the instrument are in table 1.

An Analysis of the Measure of Information Asymmetry. Since there is no measure of information asymmetry designed for use in field studies, one was developed and is reported in appendix B. The six items in the instrument are based on the definition of information asymmetry as well as on suggestions in the literature (see, e.g., Chow 1983; Chow et al. 1988; Waller and Chow 1985). The items reflect such a perspective, and the anchors were arranged on the seven-point Likert scale so that a response of (1) to any item would indicate that the superior had more information, whereas a response of (7) would indicate that the subordinate had more information than the superior. A subordinate manager had to score more than 24 on the scale for there to be any information asymmetry. A score of 24 indicates zero information asymmetry, according to the definition in this study.

Respondents' scores for information asymmetry ranged from 18 to 42. Seven managers whose scores were less than 24 were excluded from the analysis since the information differential was in favor of their superiors (i.e., there was no information asymmetry). However, the finding that only seven managers (9 percent) thought the information differential favored their superiors is consistent with the claim that subordinates typically possess local information not known to the superior (see, e.g., Baiman and Evans 1983; Chow et al. 1988).

The Cronbach alpha for the six-item instrument is 0.79. A factor analysis revealed that the items loaded on a single factor, which had an eigenvalue of 2.4086 and explained 40.1 percent of the variation in the underlying variable. The KMO measure of sampling adequacy was 0.7648. Table 1 presents descriptive statistics for this measure.

Budget Emphasis. Hopwood's (1972) eight-item measure of superiors' evaluation style was used to assess budget emphasis. Hopwood determined the styles of evaluation by the presence or absence of the items "meeting the budget" and "concern with costs" among the top three criteria of respondents. As a result of Brownell's (1982) observation that many of his respondents failed to complete the ranking properly, a seven-point

---

[3] The KMO measure assesses the extent to which the items are psychometrically related, as determined by their intercorrelations. The boundaries for the KMO are minus infinity and plus 1 (Kaiser 1970). Provided the measure of sampling adequacy is greater than 0.5, a correlation matrix is acceptable for factor-analytic examination (Kaiser and Rice 1974).

[4] Respondent interviews with eight production and six marketing managers provided anecdotal support for the findings of the psychometric analysis of the measure of budgetary slack.

Table 1
Descriptive Statistics: Independent and Dependent Variables

Variable	*Mean*	*S.D.*	*Actual Range*	*Theoretical Range*	*Cronbach Alpha*
Budgetary slack	13.528	4.806	4–25	4–28	0.6833
Budgetary participation	31.361	7.393	6–42	6–42	0.8892
Information asymmetry	31.972	5.096	24–42	24–42	0.7923
Budget emphasis	5.778	1.153	2–7	1–7	n.a.

Bivariate Correlations:

	Participation	*Information Asymmetry*
Information asymmetry	0.110	
Budget emphasis	0.147	0.061

Likert scale was used in this study for each of the eight criteria in the measure. The item "meeting the budget" alone was considered an indicator of budget emphasis. Descriptive statistics for the measure are presented in table 1.

Budgetary Participation. Subordinates' degree of participation was assessed by using the Milani (1975) six-item measure.[5] In this study, the Cronbach alpha is 0.88, which further attests to the reliability of the measure. A factor analysis revealed that all the items loaded on a single factor, which had an eigenvalue of 3.5362 and explained 58.9 percent of the variation in the underlying variable. The KMO measure of sampling adequacy was 0.8758. Descriptive statistics for the instrument are presented in table 1.

A Further Analysis of the Measures of Budgetary Slack and Information Asymmetry. To obtain further evidence of the reliability of the instruments measuring budgetary slack and information asymmetry, new questionnaires were sent to an entirely new sample of managers subsequent to hypothesis testing. Use of the same sampling criteria enabled an assessment of the stability of the measures across samples and time periods; reliability coefficients consistent with those previously obtained would show the stability of these instruments (Nunnally 1981, 191; Rosenthal and Rosnow 1991, 46).

A total of 49 additional questionnaires were mailed, of which 40 were returned. The mean age of the respondents was 43, they had held their present positions an average of four years, their mean experience in the areas they managed was 12 years, and the average number of employees in their areas of responsibility was 78. Except for the smaller number of employees managed, the descriptive statistics are similar to those for the primary sample. The reliability results also were consistent with results obtained from the primary sample. The Cronbach alphas are 0.88 for information asymmetry and 0.69 for the revised measure of budgetary slack. Since this further analysis focused only on the measures of budgetary slack and information asymmetry, the 40 responses could not be added to the primary sample.

---

[5] Evidence of the validity of the instrument has been obtained by correlating it with the Hofstede (1968) single-choice measure, for which Brownell (1983) reported a correlation of 0.74 ($p<0.01$). Brownell (1983) reported a Cronbach alpha of 0.86.

Table 2
Results of the Hypothesis Test

Variable	*Coefficient*	*Value*	*S.D.*	*t-statistics*	*p-values*
Intercept	b_0	170.400	85.350	2.00	0.050
Participation	b_1	−4.869	2.723	−1.79	0.078
Information asymmetry	b_2	−5.018	2.590	−1.94	0.057
Budget emphasis	b_3	−28.010	14.600	−1.92	0.059
Participation x information asymmetry	b_4	0.163	0.082	1.98	0.052
Participation x budget emphasis	b_5	0.862	0.459	1.88	0.065
Information asymmetry x budget emphasis	b_6	0.933	0.444	2.10	0.040
Participation x information asymmetry x budget emphasis	b_7	−0.029	0.013	−2.16	0.035

Note: $R^2=0.30$; $n=72$; $F_{7,64}=3.90$; $p=0.001$.

III. Results

The following model was used to test the hypothesis:

$$Y=b_0+b_1X_1+b_2X_2+b_3X_3+b_4X_1X_2+b_5X_1X_3+b_6X_2X_3+b_7X_1X_2X_3+e, \quad (1)$$

where:

Y=budgetary slack,
X_1=budgetary participation,
X_2=information asymmetry, and
X_3=budget emphasis.

Since it was hypothesized that the three independent variables would interact to affect budgetary slack, it was expected that the coefficient of the three-way interaction term, b_7, would be statistically significant. The results, presented in table 2, indicate that b_7 is significant at $p=0.035$ and that the model as a whole explains 30 percent of the variation in budgetary slack ($p=0.001$). The exclusion of seven managers from the analysis whose information asymmetry scores were below 24 made no difference in the results; b_7 remained significant, $t=-2.32$, $p=0.023$.[6] No attempt was made to interpret the results of significance tests for the main effects and the two-way interaction terms since the variables are not measured on ratio scales (Southwood 1978).[7]

As b_7 in table 2 is significant and negative, it appears that slack is low when

---

[6] Factor scores were also extracted with the Bartlett, regression, and Anderson-Rubin methods as provided by SPSS^X. The findings from a series of regression analyses based on the derived sets of factor scores were consistent with those using raw scores, as reported in table 2.

[7] Southwood (1978) demonstrated that the coefficients of main effects and lower-order interaction terms together with their t-statistics are arbitrary if ratio scales are not employed. He illustrated that by changing the origins of the measures of the predictors, their partial regression coefficients can be made to equal zero, since no meaningful origin can be given to non ratio-scaled data.

Table 3
Three-Way Interaction: Table of Mean Scores for Budgetary Slack

	Low Participation	*High Participation*
Low budget emphasis		
Low information: asymmetry	16.7 (6.370) (n=7)	13.1 (4.670) (n=7)
High information asymmetry	15.8 (4.981) (n=7)	16.6 (2.510) (n=5)
High budget emphasis:		
Low information asymmetry	12.0 (3.081) (n=11)	12.9 (3.885) (n=11)
High information asymmetry	15.0 (5.268) (n=9)	10.1 (4.357) (n=15)

Note: The standard deviation is given in parentheses after the mean score.

participation, information asymmetry, and budget emphasis are all high. This finding runs contrary to the expectation developed from the literature. To investigate this finding, a three-way table of slack scores based on participation, information asymmetry, and budget emphasis dichotomized at their means was constructed. Table 3, which shows the mean, standard deviation, and sample size for budgetary slack in each category, indicates that slack is lowest when all the independent variables are high, and highest when all the predictors are low.

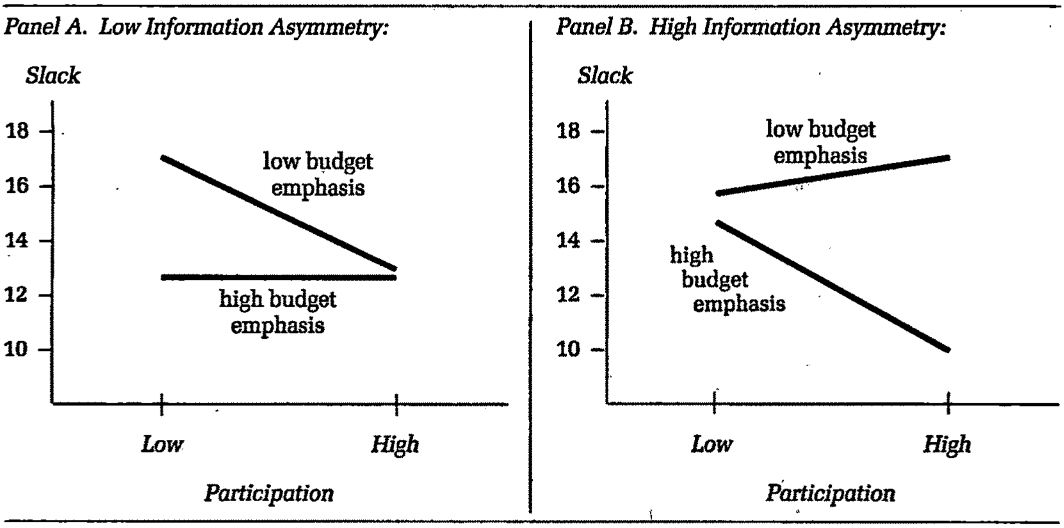
The information in table 3 was used to graph the three-way interaction shown in figure 1. The two graphs illustrate that slack is lowest when information asymmetry, participation, and budget emphasis are all high, and that slack is highest when the predictors are all low. The graphs also illustrate that slack is higher when budget emphasis is low rather than high.

IV. Conclusions

The results of this study show that the relation between participation and slack is contingent upon budget emphasis and information asymmetry, but in a direction contrary to expectations. The results provide evidence for the utility of participative budgeting, and little support for the view that high participation may result in increased slack when the other two predictors are high. Although participation may induce subordinates to incorporate slack in budgets, the results suggest that participation alone may not be sufficient. The findings suggest that slack reduction results from participation, except when budget emphasis is low.

The results of the study may be subject to a number of possible limitations. Measures of two constructs had to be developed, and their validation has not been independently established. Also, construction of the measures of budgetary slack and information asymmetry from the literature reviews, item analyses, and factor analyses may have driven the results. In addition, it could be argued that the findings obtained in this research are partly attributable to a poor measure of the dependent variable

Figure 1
Three-Way Interaction Graphs


Panel A. Low Information Asymmetry:

Panel B. High Information Asymmetry:

(budgetary slack). As a construct, slack may be seen in both a clinical and an emotive sense. Although the interest here is in the former concept, managers' questionnaire responses may have been founded on an emotive platform and the results could be prejudiced.

Another limitation of the study is that data were drawn only from firms located in the Sydney metropolitan area; hence, the results are potentially only generalizable to that population. It may also be the case that cross-sectional survey methods are not sufficient for the investigation of budgetary slack, given the nature of the construct. Case studies may assist in providing additional information on the circumstances in which slack may occur.

Appendix A
Budgetary Slack Measure

The following statements relate to the budgetary environment in which you work. Please indicate the extent of your agreement with each statement by circling a number from 1 to 7, based on the following scale:

1. Strongly disagree
2. Moderately disagree
3. Mildly disagree
4. Neutral
5. Mildly agree
6. Moderately agree
7. Strongly agree

1. Standards set in the budget induce high productivity in my area of responsibility.*	1	2	3	4	5	6	7
2. Budgets set for my area of responsibility are safely attainable.	1	2	3	4	5	6	7
3. I have to carefully monitor costs in my area of responsibility because of budgetary constraints.*	1	2	3	4	5	6	7
4. Budgets for my area of responsibility are not particularly demanding.	1	2	3	4	5	6	7
5. Budgetary targets have not caused me to be particularly concerned with improving efficiency in my area of responsibility.	1	2	3	4	5	6	7
6. Targets incorporated in the budget are difficult to reach.*	1	2	3	4	5	6	7

*Responses to these items were reverse-scored.

Appendix B
Information Asymmetry Measure

Please respond to each of the following questions by circling a number from 1 to 7, based on their individual scales:

1. In comparison with your superior, who is in possession of better information regarding the activities undertaken in your area of responsibility?

1	2	3	4	5	6	7
My superior has much better information			We have about the same quality of information			I have much better information

2. In comparison with your superior, who is more familiar with the input-output relationships inherent in the internal operations of your area of responsibility?

1	2	3	4	5	6	7
My superior is much more familiar			We are about equally familiar			I am much more familiar

3. In comparison with your superior, who is more certain of the performance potential of your area of responsibility?

1	2	3	4	5	6	7
My superior is much more certain			We are about equally certain			I am much more certain

Continued

Appendix B—*Continued*

4. In comparison with your superior, who is more familiar technically with the work of your area of responsibility?

1	2	3	4	5	6	7
My superior is much more familiar			We are about equally familiar			I am much more familiar

5. In comparison with your superior, who is better able to assess the potential impact on your activities of factors external to your area of responsibility?

1	2	3	4	5	6	7
My superior is much better able			We are about equally able			I am much better able

6. In comparison with your superior, who has a better understanding of what can be achieved in your area of responsibility?

1	2	3	4	5	6	7
My superior has a much better understanding			We have about the same understanding			I have a much better understanding

References

Antle, R., and G. D. Eppen. 1985. Capital rationing and organizational slack in capital budgeting. *Management Science* 31 (February): 163–74.

Baiman, S. 1982. Agency research in managerial accounting: A survey. *Journal of Accounting Literature* 1 (Spring): 154–213.

——, and J. H. Evans III. 1983. Pre-decision information and participative management control systems. *Journal of Accounting Research* 21 (Autumn): 371–95.

——, and B. Lewis. 1989. An experiment testing the behavioral equivalence of strategically equivalent employment contracts. *Journal of Accounting Research* 27 (Spring): 1–20.

Brownell, P. 1982. The role of accounting data in performance evaluation, budgetary participation, and organizational effectiveness. *Journal of Accounting Research* 20 (Spring): 12–27.

——. 1983. Leadership style, budgetary participation and managerial behavior. *Accounting, Organizations and Society* 8 (4): 307–21.

Cammann, C. 1976. Effects of the use of control systems. *Accounting, Organizations and Society* 1 (4): 301–13.

Chow, C. W. 1983. Providing incentives to limit budgetary slack. *Cost and Management* 57 (September-October): 37–41.

——, J. C. Cooper, and W. S. Waller. 1988. Participative budgeting: Effects of a truth-inducing pay scheme and information asymmetry on slack and performance. *The Accounting Review* 63 (January): 111–22.

Christensen, J. 1982. The determination of performance standards and participation. *Journal of Accounting Research* 20 (Autumn): 589–603.

Coughlan, A. T., and R. M. Schmidt. 1985. Executive compensation, management turnover, and firm performance: An empirical investigation. *Journal of Accounting & Economics* 7: 43–66.

Dominowski, R. L. 1980. *Research Methods*. Englewood Cliffs, NJ: Prentice-Hall.

Hofstede, G. H. 1968. *The Game of Budget Control*. Assen, The Netherlands: Van Gorcum.

Hopwood, A. G. 1972. An empirical study of the role of accounting data in performance evaluation. *Journal of Accounting Research* 10 (Supplement): 156–82.

Kaiser, H. F. 1970. A second generation little jiffy. *Psychometrika* 35 (December): 401–15.

——, and J. Rice. 1974. Little jiffy, mark IV. *Educational and Psychological Measurement* 34: 111–17.

Kompass Australia. 1988. Prahan, Australia: Peter Isaacson.

Kren, L., and W. M. Liao. 1988. The role of accounting information in the control of organizations: A review of the evidence. *Journal of Accounting Literature* 7: 280–309.

Lowe, E. A., and R. W. Shaw. 1968. An analysis of managerial biasing: Evidence from a company's budgeting process. *The Journal of Management Studies* 5 (October): 304–15.

Lukka, K. 1988. Budgetary biasing in organizations: Theoretical framework and empirical evidence. *Accounting, Organizations and Society* 13 (3): 281–301.

Magee, R. P. 1980. Equilibria in budget participation. *Journal of Accounting Research* 18 (Autumn): 551–73.

Merchant, K. A. 1985a. Budgeting and the propensity to create budgetary slack. *Accounting, Organizations and Society* 10 (2): 201–10.

———. 1985b. *Control in Business Organizations*. Marshfield, MA: Pitman.

Milani, K. 1975. The relationship of participation in budget-setting to industrial supervisor performance and attitudes: A field study. *The Accounting Review* 50 (April): 274–84.

Monette, D. R., T. J. Sullivan, and C. R. DeJong. 1986. *Applied Social Research*. New York: Holt, Rinehart and Winston.

Nunnally, J. C. 1981. *Psychometric Theory*. New Delhi, India: Tata McGraw-Hill.

Onsi, M. 1973. Factor analysis of behavioral variables affecting budgetary slack. *The Accounting Review* 48 (July): 535–48.

Penno, M. 1984. Asymmetry of pre-decision information and managerial accounting. *Journal of Accounting Research* 22 (Spring): 177–91.

Pope, P. F. 1984. Information asymmetries in participative budgeting: A bargaining approach. *Journal of Business Finance & Accounting* 11 (Spring): 41–59.

Rosenthal, R., and R. L. Rosnow. 1991. *Essentials of Behavioral Research: Methods and Data Analysis*. New York: McGraw-Hill.

Schiff, M., and A. Y. Lewin. 1968. Where traditional budgeting fails. *Financial Executive* 36 (May): 50–62.

———, and ———. 1970. The impact of people on budgets. *The Accounting Review* 45 (April): 259–68.

Southwood, K. E. 1978. Substantive theory and statistical interaction: Five models. *American Journal of Sociology* 83 (March): 1154–1203.

Waller, W. S. 1988. Slack in participative budgeting: The joint effect of a truth-inducing pay scheme and risk preferences. *Accounting, Organizations and Society* 13 (1): 87–98.

———, and C. W. Chow. 1985. The self-selection and effort effects of standard-based employment contracts: A framework and some empirical evidence. *The Accounting Review* 60 (July): 458–76.

Williamson, O. E. 1964. *The Economics of Discretionary Behavior: Managerial Objectives in a Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.

Young, S. M. 1985. Participative budgeting: The effects of risk aversion and asymmetric information on budgetary slack. *Journal of Accounting Research* 23 (Autumn): 829–42.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 411–416

COMMENTS

A Comment on "A Multidimensional Analysis of Selected Ethical Issues in Accounting"

Scott K. Jones
University of Delaware
Lawrence A. Ponemon
Babson College and SUNY Binghamton

IN a recent article, Flory et al. (1992) reported on the reliability, content validity, and predictive validity of a multivariate ethical measurement scale. Their motivation was to investigate factors that "might influence ethical behavior, realizing that these factors are potentially complex and multifaceted" (p. 288). Their overall conclusion was that the scale was highly reliable and valid, and thus should be a "potentially useful tool for directing research on why accountants make certain ethical judgments" (p. 298). We congratulate the authors for undertaking an important area of empirical research, but believe that the multidimensional scale fails to provide a useful psychometric construct for measuring the ethical thinking process of professional accountants. Furthermore, though the rigor of their tests of validity is a commendable model for all researchers, the reliability achieved should not be a gauge for those scales reported or for any other similarly intentioned scales.

Our criticism of Flory et al. (1992) is twofold. First, by failing to incorporate or expressly consider a psychological framework for the ethical reasoning process, the multidimensional ethics scales developed by Flory et al. do not provide any insight into the fundamental question *What makes accountants more or less ethical*? Second, by the very nature of the research design, the otherwise well-chosen measures of reliability and validity are biased toward high measurement scale reliability.

I. Correspondence to Psychological Theory

The psychology of ethical reasoning provides a theory of the process that individuals follow in making a decision. On the basis of earlier work in this area by Piaget

Submitted May 1992.
Accepted October 1992.

Table 1
Kohlberg's Six Stages of Ethical Reasoning

Stage	*What Is Right?*	*Reason For Doing It*
Pre-Conventional Level:		
Stage 1	Avoid breaking rules backed by punishment	Avoidance of punishment or harm to self
Stage 2	Follow rules if they are in own interest	To serve one's own immediate interests
Conventional Level:		
Stage 3	Living up to what is expected by others	Need to be a good person in the eyes of others
Stage 4	Fulfilling agreed-to duties and obligations	To keep the social institutions going
Post-Conventional Level:		
Stage 5	Upholding non-relative obligations first	Obligation to law before social contracts
Stage 6	Follow self-chosen ethical principles	Belief in ideal as a rational person

Adapted from Kohlberg (1969).

(1932), Kohlberg (1969) developed a theory of moral development, advancing a stage-sequencing model somewhat akin to the rungs of a ladder (see table 1). According to Kohlberg, starting at birth, all individuals progress upwardly through a series of cognitive equilibria, from what is termed "pre-conventional morality" to "conventional morality" and sometimes to the final and highest level, "post-conventional morality" (see Kohlberg 1984, 624–39).

Using this stage-sequencing model, researchers have constructed reliable psychometric instruments to assess an individual's level of ethical reasoning. Two well-known and widely used measures are the moral judgment interview (Colby and Kohlberg 1987) and the defining issues test (DIT) of Rest (1979). Rest (1986) summarizes the findings of more than 500 studies using the DIT, which provides a highly reliable and externally valid measure of post-conventional ethical reasoning. Drawing on the work of Kohlberg (1969) and Rest (1979), numerous studies have attempted to examine accountants' levels of ethical reasoning and moral development within the context of judgment and decision making (see, e.g., Armstrong 1987; Arnold and Ponemon 1991; Hiltebeitel and Jones 1991, 1992; Lampe and Finn 1992; Ponemon 1990, 1992a, 1992b; Ponemon and Gabhart 1992; Ponemon and Glazer 1990; St. Pierre et al. 1990). In general, these studies report that accountants and accounting students are not developing ethical reasoning capacities commensurate with individuals of similar socioeconomic or educational backgrounds. More importantly, these studies corroborate the importance of ethical reasoning as a determinant of judgment, behavior, and action in the practice of accounting and auditing.

Although the cognitive-developmental theories espoused by Kohlberg (1969) and Rest (1979) apply to a broad class of social and ethical conflicts, including those defined in the Flory et al. article, no correspondence with, or mention of, related works in this area is provided by Flory et al. In the absence of a large-scale factorial design, the multidimensional ethics scales developed and used by Flory et al. cannot address the research question they pose since they treat the cognitive process as a proverbial "black box." Without experimenting with exhaustive combinations of relevant input/output variables, the reader can only imagine how accountants think about, frame, or resolve ethical issues. In this regard, the psychology of ethical reasoning has repeatedly been shown to be of consequence to behavior in the context of accounting and auditing because many professional judgments are conditioned on the beliefs and values of the individual practitioner.

In light of prior reseach, Flory et al. should have acknowledged the existence of relevant work in the ethics field and developed correspondence between their scales and validated psychometric measures of ethical reasoning (or, at a minimum, justify the need for alternative, largely untested, instrumentation). It is surprising that this was not done, given that Reidenbach and Robin (1990, 649), two of the authors in Flory et al., state in an earlier work that "the three positive dimensions do not correspond strictly to the normative moral philosophies and tend to disagree with several of the hypothesized relationships in recently developed models of ethical decision-making." Reidenbach and Robin go on to suggest that the multidimensional ethics scales developed by them could enhance the predictability of Kohlberg's and Rest's works but fail to establish any correspondence with those theories. Had Flory et al. done so, their conclusions may have been less ambiguous and may have suggested new directions for future ethics research.

II. Scale Reliability

Regardless of correspondence with extant theories of ethical reasoning, a significant methodological problem with Flory et al. is that the realistic reliability of their multidimensional ethics scales is likely lower than reported because of a serious sampling bias that is attributable to three flaws in their research design:

1. The subjects are not randomly sampled neither among accountants nor members of the Institute of Management Accountants (IMA), and thus are far more homogeneous than the population of accounting professionals as a whole;
2. The IMA/Ethics Resource Center cases are well known and represent situations of clearly defined ethical conflict; and
3. The resolving action in each hypothetical case produces a contrast effect with the prescribed IMA code of conduct.

The homogeneity of subjects and the clear conflict in the cases result in sampling bias because commonly encountered situations would involve more diversity and the ethical conflict would typically be less clear. For example, CMAs may have prior knowledge of the IMA cases and likely represent a more homogeneous profile in their responses than non-CMAs. Furthermore, inasmuch as the issues in the chosen cases are easily identifiable by their familiarity, realistic ethical dilemmas are not represented. As noted by Kidder (1981, 132), "If measures are so similar that they are guaranteed to produce 100 percent agreement, they would not be considered a fair test of the reliability of an instrument." Though the design chosen by Flory et al. does not guarantee agreement, it goes a long way toward assuring very little disagreement.

Reliability is one aspect of response scale measurement that can have serious consequences for internal validity.[1] In Flory et al. there exists a contrast effect bias that systematically overstates reported reliability. Contrast effect refers to the tendency of raters to see others' traits and actions the opposite of their own (Kidder 1981, 206). The idea is based on social judgment theory (Sherif and Hovland 1964). Basically, when an attitudinal statement such as the actions chosen by Flory et al. is within a subject's bounds of rejection, the attitude represented is rejected with more vigor than is warranted by the true position of the subject. The reaction is most pronounced when an individual is actively involved, as would be the case for CMAs faced with ethical conflict. Similarly, even if subjects agree with an action (i.e., it is within their bounds of acceptance), their homogeneity would lead to an assimilation effect, whereby they see the action more favorably than they would normally. Since the scenarios are presented in a role-playing context, social bias is not remedied by the use of a third-person construct.

Suffice to say that the exclusive use of CMAs and four of the five well-known IMA/Ethics Resource Center cases must raise questions about whether the scales are as reliable and valid as would be the case had the resolving action been different. All four actions specified in the experiment could easily be interpreted as violations of prescribed *guidelines* for the resolution of ethical conflict, aside from the desirability of the action specified. The guidelines of *Statements on Management Accounting No. 1C* (SMA No. 1C, IMA 1983) for resolving ethical conflict clearly specify that the management accountant is to look to established company policy, and if policy does not resolve that conflict, the management accountant is to adhere to the following sequence until the conflict is resolved.:

1. Discuss the problem with immediate supervisors or the next higher managerial level, if necessary;
2. Clarify in confidential discussions with an advisor; or
3. Resign from the organization.

The actions specified by Flory et al. are essentially equivalent to turning one's back on the issue, and there is no indication that the resolution guidelines were adhered to, which itself is clearly in violation of SMA No. 1C and prone to a contrast effect as described above.

Since all of the subjects were CMAs, and the actions were essentially contrary to what would be implied by SMA No. 1C, the subjects were much more disposed to viewing themselves as ethical and the specified action as unethical. The resulting systematic effect erroneously increases the reliability of the ratings because the subjects' positions would agree, and they would rate the actions by converging on extremes of the scales. The multivariate, univariate, and behavioral intention means presented in table 4 of Flory et al. (1992, 295) suggest an overwhelming convergence toward the "unfair, unjust, and unacceptable" side of the scales. The overwhelming agreement is evidenced by the domination of means well above the midpoints of the scales. Flory et al. suggest that end-loading is absent (p. 295), but since no standard deviations are reported, since means above the midpoints dominate, and since there is an *a priori* argument supporting contrast effect, it is reasonable to infer that the reported means are statistically different from the midpoints and that in fact there is substantial end-loading on all scales in the same direction.

---

[1] For a full discussion of the threats to validity of these and other factors, see Cook and Campbell (1979, 51) and Campbell and Stanley (1963, 5).

Because the chosen measures of reliability and validity are related to correlation, the statistical effect of convergence attributable to a contrast effect is sufficiently illustrated in the formulation of the alpha coefficient. Cronbach's alpha, one measure of reliability, is the squared correlation between the observed (X_i) and true (t_i) scores. If the measurement error for the observed score of scale item (i) is independent of the true score, Cronbach's alpha can be computed as follows:

$$p^2(X_i,t_i)=\frac{Var(t_i)}{Var(X_i)}, \tag{1}$$

where Ω is a random variable such that:

$$t_i=\mu_i+\beta_i\Omega, \tag{2}$$

$$X_i=\mu_i+\beta_i\Omega+e_i. \tag{3}$$

It follows that a contrast effect would increase reliability because X_i converges to a ceiling or floor value on the scale as the subjects converge on one end of the scale, given that the actions are all within the bounds of rejection. Similarly, by equation (3), the variance of the observed scores (X_i) decreases as the e_i converge to zero, which reduces the random error used in all of the univariate and multivariate statistics. Since contrast would have the effect of limiting the variation in e_i, the scales will appear more reliable and valid than if not subjected to this design problem. It can also be argued that, in the underestimation of the measurement error (e_i), the reported beta coefficients and adjusted R^2 values are overstated both in magnitude and significance. More realistic beta coefficients and adjusted R^2 values would be attenuated by the bias induced from measurement error (Fuller 1987, 5).

Although the aforementioned problems may not invalidate the ethics scales used by Flory et al., these scales and the reliability purportedly represented should not be the standard by which subsequent ethics research is gauged. In the absence of a more randomized design, the high reliability and validity reported must be viewed with extreme skepticism. To generalize the reliability of these scales, it would be necessary to select both the subjects and cases randomly and thereby reduce the bias from contrast or assimilation effects with the resolving action. The reliability and validity tests chosen are not disputed, but the realistic reliability of these scales is likely less—and perhaps much less—than reported levels.

To beg semantics, the answer to their research question is obvious—accountants make ethical decisions because they must do so to avoid violating their professional responsibilities. The more relevant empirical question is how accountants make judgments in ethical situations, and this remains largely unanswered. The multidimensional scales suggested by Flory et al. are inadequate for an understanding of the process accountants use to resolve ethical conflicts because they are based solely on a biased design. Researchers should be expected to perform tests of reliability and validity, and at least Flory et al. can be a model for that inquiry.

References

Armstrong, M. 1987. Moral development and accounting education. *Journal of Accounting Education* 5 (Spring): 27–43.

Arnold, D., and L. Ponemon. 1991. Internal auditors' perceptions of whistle-blowing and the influence of moral reasoning: An experiment. *Auditing: A Journal of Practice & Theory* 10 (Fall): 1–15.

Campbell, D. T., and J. C. Stanley. 1963. *Experimental and Quasi-experimental Designs for Research*. Boston: Houghton Mifflin.

Colby, A., and L. Kohlberg. 1987. *The Measurement of Moral Judgment*. New York: Cambridge University Press.

Cook, T. D., and D. T. Campbell. 1979. *Quasi-experimentation—Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin.

Flory, S. M., T. J. Phillips, R. E. Reidenbach, and D. P. Robin. 1992. A multidimensional analysis of selected ethical issues in accounting. *The Accounting Review* 67 (April): 284–302.

Fuller, W. A. 1987. *Measurement Error Models*. New York: John Wiley & Sons.

Hiltebeitel, K., and S. Jones. 1991. Initial evidence on the impact of ethics integration in accounting education. *Issues in Accounting Education* 6 (Fall): 262–75.

——, and ——. 1992. An assessment of ethics instruction in accounting education. *Journal of Business Ethics* 11: 37–46.

Institute of Management Accountants. 1983. *Statements on Management Accounting No. 1C—Standards of Ethical Conduct for Management Accountants*. Montvale, NJ: IMA.

Kidder, L. H. 1981. *Research Methods in Social Relations*. New York: Holt, Rinehart and Winston.

Kohlberg, L. 1969. Stages and sequences: The cognitive developmental approach to socialization. In *Handbook of Socialization Theory and Research*, edited by D. Goslin. Chicago: Rand McNally.

——. 1984. *Essays on Moral Development*. Vol. 2. *The Psychology of Moral Development*. San Francisco: Harper & Row.

Lampe, J., and D. Finn. 1992. A model of auditors' ethical decision process. *Auditing: A Journal of Practice & Theory*. In press.

Piaget, J. 1932. *The Moral Judgment of the Child*. Reprint, New York: Free Press, 1966.

Ponemon, L. 1990. Ethical judgments in accounting: A cognitive-developmental perspective. *Critical Perspectives in Accounting* 1(2): 191–215.

——. 1992a. Ethical reasoning and selection socialization in accounting. *Accounting, Organizations and Society* 17 (April/May): 239–58.

——. 1992b. Auditor underreporting of time and moral reasoning: An experimental-lab study. *Contemporary Accounting Research*. In press.

——, and D. Gabhart. 1990. Auditor independence judgments: A cognitive-developmental model and experimental evidence. *Contemporary Accounting Research* 7 (Fall): 227–51.

——, and A. Glazer. 1990. Accounting education and ethical development: The influence of liberal learning on students and alumni in practice. *Issues in Accounting Education* 5 (Fall): 21–34.

Reidenbach, R., and D. Robin. 1990. Toward the development of a multidimensional scale for improving evaluations of business ethics. *Journal of Business Ethics* 9: 639–53.

Rest, J. 1979. *Development in Judging Moral Issues*. Minneapolis: University of Minnesota Press.

——. 1986. *Moral Development: Advances in Research and Theory*. New York: Praeger.

Sherif, M., and C. Hovland. 1964. *Social Judgment*. New Haven, CT: Yale University Press.

St. Pierre, E. K., E. Nelson, and A. Gabbin. 1990. A study of the ethical development of accounting majors in relation to other business and non-business disciplines. *The Accounting Educators Journal* 2 (Summer): 23–35.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 417–421

A Reply to "A Comment on 'A Multidimensional Analysis of Selected Ethical Issues in Accounting'"

Steven M. Flory
Thomas J. Phillips, Jr.
Louisiana Tech University
R. Eric Reidenbach
Donald P. Robin
University of Southern Mississippi

WE welcome the opportunity to respond to Jones and Ponemon's (1993) comment on our article "A Multidimensional Analysis of Selected Ethical Issues in Accounting." The authors of the comment focus on two issues that will be addressed in the same order as presented by them.

The first issue concerns the failure to include in our study the work of Kohlberg and Rest. In response, we will indicate how the multivariate analysis in our approach is theoretically different from the moral development work of Kohlberg and Rest and, further, how it is potentially a more productive means for understanding the ethical decision processes that are related to specific situations in accounting.

The second issue involves the high reliability measures that were present in our research. The authors of the comment speculate about factors that might have produced inordinately high coefficient alphas. We explain why these concerns are unfounded and suggest that future researchers using the multivariate approach can expect similarly high results.

I. The Relationship of the Multidimensional Approach to the Moral Development Construct

As the authors of the comment note, some work in the accounting literature uses the defining issues test (DIT) by Rest (1979, 1986) for measuring moral development in accounting. Although this work on moral development represents an important stream of research, it is substantially different from the approach suggested by our research.

Invited by the Editor.
Accepted November 1992.

Further, these differences are extremely important if the objective of ethics research is to explain, predict, and control unethical behavior by accountants.

The multivariate approach is situation-specific, temporal, and process-oriented, while the moral development approach is general (not situation-specific), enduring, and trait-oriented (see table 1). The DIT uses scenarios unrelated to ethical problems in accounting to place individuals in one of the six stages of moral development according to the relative importance of different justifications by respondents. The expectation is that individuals who use a certain method of reasoning for the DIT scenarios will continue to use the same approach for ethical reasoning in other situations. Although people may develop morally over time, the process is slow, and an individual is likely to remain in a single stage for many years. For that reason, we have labeled an individual's stage of moral development as a generalized enduring trait, like personality. Unfortunately, like the personality constructs, moral involvement has not exhibited strong relationships with behavioral measures (Rest 1979, 1986).

The multivariate approach described in Flory et al. (1992) describes situations that are specific and relevant to accounting. Our study used more elaborately specified descriptions of ethically questionable situations than any of the previous development work on the measure. The multivariate approach recognizes that specific subtleties in an ethical situation can change both the ethical judgment and behavior of an accountant. It also recognizes that external and internal pressures on the accountant constantly change and can produce different results in judgment and behavior. For these reasons, we believe that studying the ethical decision process is likely to be a more productive approach for explaining, predicting, and controlling the ethical behavior of accountants.

The last entry in table 1 provides some empirical justification for our beliefs about the productivity of the multivariate approach. Empirical results with Rest's DIT have generally exhibited poor predictive validity (i.e., the results have little correlation with behavioral intent). Rest (1979, 259–61) recognizes these limitations and attempts to justify the weak predictive results. In a later book, Rest (1986, 136–41) presents a table of "Studies relating DIT scores to a measure of behavior" (table 5.1) and provides correlations and sample sizes for 13 studies. Changing these correlations to adjusted R^2s for direct comparison with the findings in our article produced the following results. Three of the studies cited by Rest produced nonsignificant results, with adjusted R^2s that were essentially zero. The adjusted R^2s in the other studies included three that were less than 0.05, three that were between 0.05 and 0.10, three that were between 0.10 and 0.17, and one that was 0.22. The adjusted R^2s for predicting behavioral intent in Flory et al. (1992) ranged from 0.45 to 0.76. Further, other research to develop and apply the multivariate measure has produced consistently higher adjusted R^2s than those cited above for the DIT (see Humphreys et al., forthcoming; Reidenbach and Robin 1990; Reidenbach et al. 1991).

The authors of the comment state "the multidimensional scale fails to provide a useful psychometric construct for measuring the ethical thinking process of professional accountants" and that "the multidimensional ethics scale developed by Flory et al. do not provide any insight into the fundamental question *What makes accountants more or less ethical?*" Clearly, both statements are untrue if the explanation, prediction, and control of unethical behavior is the objective.

The differences between the multidimensional and moral development approaches

Table 1
A Comparison of the Moral Development and Multidimensional Approaches for Understanding Ethical Behavior in Accounting

Criteria for Comparison	*Moral-Development Characteristics*	*Characteristics of the Multidimensional Approach of Flory et al. (1992)*
Specificity	Generalized/nonspecific focus	Situational/specific focus
Time	Enduring (results apply over blocks of time)	Temporal (results can change over time)
Orientation for explaining behavior	Trait orientation	Process orientation
Ability to predict ethical or unethical behavior	Poor (majority of adjusted R^2s have been less than 0.10)	Acceptable to good (adjusted R^2s between 0.45 and 0.76 in Flory et al.)

identified in table 1 explain why the work on moral development was excluded from Flory et al. (1992). Journals correctly offer limited space for background information, and the work on moral development was considered to be different enough from the multidimensional approach to warrant its omission.

II. Scale Reliability

The comments and concerns of the authors listed under this heading are puzzling. The stated objective of our original article was that of construct development, and a systematic sample of ICMA members was used to aid that objective. It was never our intent to speak to "the population of accounting professionals as a whole." In research, it is perfectly acceptable to select a population of interest and work within it. Perhaps more important, unless all accounting professionals are believed to behave alike (in which case the comment on Flory et al. would have no meaning), it would be scientifically interesting to compare and contrast different groups of accounting professionals to the ICMA members used in our study. It seems obvious to point out that the objective of such studies is to minimize the within-group variance in comparison with between-group variance.

The concern that "the IMA/Ethics Resource Center cases are well known and represent situations of clearly defined ethical conflict" is also puzzling. First, the reactions of accountants to "clearly defined ethical conflict" is exactly what was intended in the study. Perhaps a hint regarding this concern comes from the later statement that "the answer to their research question is obvious—accountants make ethical decisions because they must do so to avoid violating their professional responsibilities." Are we to assume that accountants are automatons with no prior understanding or training in ethics, and completely subservient to their codes of ethics? The naivety of such a statement suggests a superficial understanding about the ethical behavior of accountants. It confuses *normative* statements about what accountants ought to do with *positive* statements about their actual behavior. Our respondents had a much wider diversity of beliefs than the authors of this comment seem to understand.

The distribution of each of our measures extended the full limits of the scale (1 to 7), and the standard deviations ranged from 1.1 to 1.7—more than one scale point for each of the four scenarios.

The final concern in the comment is with the existence of a contrast effect. There is absolutely no evidence that a contrast effect is operating in Flory et al. (1992). It is pure speculation to suggest that such an effect exists. If the authors are firm in their belief about this hypothetical problem, they should undertake the empirical research necessary to determine if the contrast effect is real. However, there are reasons to believe that such research would be fruitless.

The most important response to concerns about the high reliability coefficients found in our study comes from corroborating research. If this had been the first and only study using the multivariate construct, then questions about these statistical features would be of more concern. However, additional research discussed and referenced in the article has also produced good reliability coefficients. For example, in Reidenbach and Robin (1990), which was cited in Flory et al. (1992), coefficient alphas ranged from 0.71 to 0.92 and averaged 0.80. These results compare favorably with the range of 0.75 to 0.94 and average of 0.86 reported in Flory et al., and other published findings are similar. Further, these studies used widely different scenarios as well as different subjects, and there is little reason to believe that future efforts will not be similarly effective (see Humphreys et al., forthcoming; Reidenbach and Robin 1990; Reidenbach et al. 1991).

III. Conclusion

As we stated in Flory et al. (1992, 228), "Psychometricians almost never regard measure completely valid in every circumstance, but rather consider the process of validation to be unending. Further, measures of potentially broad constructs, such as 'ethics,' must be tested in a variety of settings and with different groups to confirm that the initial results are not an artifact of a particular group or setting." Our research represents an initial application of this measure to accounting, but much additional research is needed in the discipline to validate the construct. We invite all others who might be interested to join us in this continuing validation effort.

References

Flory, S., T. Phillips, E. Reidenbach, and D. Robin. 1992. A multidimensional analysis of selected ethical issues in accounting. *The Accounting Review* 67 (April): 284–302.

Humphreys, N., D. Robin, E. Reidenbach, and D. Moak. Forthcoming. The ethical decision-making process of small business owner/managers and their customers. *Journal of Small Business Management.*

Jones, S. K., and L. A. Ponemon. 1993. A comment on "A multidimensional analysis of selected ethical issues in accounting." *The Accounting Review* 68 (April): 411–16.

Kohlberg, L. 1969. Stages and sequences: The cognitive developmental approach to socialization. In *Handbook of Socialization Theory and Research*, edited by D. Goslin. Chicago: Rand McNally.

Reidenbach, E., and D. Robin. 1990. Toward the development of a multidimensional scale for improving evaluations of business ethics. *Journal of Business Ethics* 9 (August): 639–53.

———, ———, and L. Dawson. 1991. An application and extension of a multidimensional ethics scale to selected marketing practices and marketing groups. *Journal of the Academy of Marketing Science* 19 (Spring): 83–92.

Rest, J. 1979. *Development in Judging Moral Issues*. Minneapolis: University of Minnesota Press.

———. 1986. *Moral Development: Advances in Research and Theory*. New York: Praeger.

THE ACCOUNTING REVIEW
Vol. 68, No. 2
April 1993
pp. 422–438

BOOK REVIEWS

Shane R. Moriarity, Editor

Editor's Note: Two copies of books for review should be sent to Professor Shane Moriarity, School of Accounting, University of Oklahoma, Norman, OK 73019. The policy of the *Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

D. M. BARNES, *Value-for-Money Audit Evidence* (Toronto: The Canadian Institute of Chartered Accountants, 1992, pp. xii, 117, C$35.00, paper).

The Canadian Institute of Chartered Accountants commissioned D. M. Barnes, principal author of this Research Study, to consider the concept of audit evidence as it relates to value-for-money (VFM) audits. VFM audits focus on determining whether organizations and/or programs are managed in an economical, efficient, and effective manner. Although VFM audits have found a niche in the public sector, VFM auditing can also be applied to assess the efficiency and effectiveness of organizational units in the private sector. The purpose of this study is to stimulate discussion and debate concerning the nature, gathering, and evaluation of audit evidence for VFM audits.

Early sections of this study devote considerable attention to defining basic terms such as "fact," "finding," and "conclusion." Although discussion of key terms such as "sufficiency" and "appropriateness" are clearly warranted, I found this section of the study to be somewhat pedantic.

Both VFM audits and financial statement audits are governed by professional standards requiring that "sufficient appropriate audit evidence should be obtained..." (pp. 6–7). However, the author contends that since the nature and circumstances of VFM audits vary widely from financial statement audits, the concepts of audit evidence will also differ. For example, oral evidence from interviews plays a far more critical and significant role in VFM audits than it does in financial statement audits. In addition, mathematical probability assumes a less significant role in assessing the sufficiency and appropriateness of evidence in VFM audits than it does for a financial statement audit.

Barnes relied on two primary sources of information in developing his positions. First, he conducted a review of Canadian and International literatures on audit evidence, legal evidence, and logic and reasoning. The second primary source of information was gathered from a survey sent to a cross-section of practitioners concerned with VFM audits. Supplemental interviews were conducted with individuals knowledgeable in different areas of evidence.

The applicability of three alternative conceptual models was assessed to provide a basic framework for VFM audit evidence. The relevance of financial statement audit evidence and legal evidence concepts was evaluated. The study also examined logic and reasoning concepts as they relate to audit evidence.

A useful contribution of this book is the development of a VFM audit evidence logic model. Logic theory provides the underlying framework for this model. "The model highlights the critical decision points in a VFM audit... (and)... displays the reasoning process of the audit superimposed on the stages of the audit process" (p. 37). I believe that this model should help readers focus on the sufficiency and appropriateness of evidence for VFM audits.

The study considered evidential concepts in a financial statement audit and concluded that the emphasis is predominantly quantitative. However, VFM audits are concerned with measures of management performance and are not constrained to monetary measures. Since VFM audits emphasize qualitative attributes, the author rejected the quantitative approach underlying evidence accumulation in a financial statement audit as a basis for their model.

The author also considered how legal systems provide or support models of evidence. He acknowledged some parallels between legal and VFM audit evidence. However, fundamental differences between the concepts of legal and VFM audit evidence, such as the adversarial role of the legal process, precluded reliance on a legal model.

Inclusion of practitioner views is another important contribution of this book. They add a dimension of pragmatism to an otherwise theoretical treatise of the concept of audit evidence. The author notes that "The views expressed are not intended to be a statistically valid sample of current practice—instead they are a sampling of opinions from individual practitioners" (p. 65). Some practitioner responses concerned the author. For example, many respondents indicated that "if the auditee does not object to a finding, then no further evidence is required" (p. 69). The author argues that all findings should be supported by objective persuasive evidence.

In summary, this study should interest all concerned with VFM or performance audits. For readers unfamiliar with VFM audits and/or basic concepts of audit evidence, the early chapters provide a thorough discussion of evidential concepts and explain several differences in VFM and financial statement audits. Practitioners should find this book enlightening as it focuses on issues pertaining to the collection and evaluation of audit evidence. They will find Chapters 4 through 10 of particular interest. These chapters present the development of the VFM audit logic model and discuss its relationship to obtaining "sufficient appropriate audit evidence." Practice issues obtained from surveys are also evaluated. Researchers will also find interesting avenues for further inquiry in an area that has received relatively little attention.

PAUL A. JANELL
Professor of Accounting

Northeastern University

VINOD B. BAVISHI, *Global Company Handbook: An Analysis of the Financial Performance of the World's Leading 7,500 Companies, 1992 Edition* (Princeton: Center for International Financial Analysis & Research, Inc., 1992, pp. 2,500 in 2 volumes, $395.00, paper).

This two-volume publication is a global database containing limited financial and operating information for more than 7,500 of the world's largest and most successful companies. Included in the database are firms from 48 countries in 39 industry classes. This resource is intended to be used in applications which require comparative information for global enterprises. Accordingly, certain nominal financial data are denoted in U.S. dollars for all firms in the database. Additionally, CIFAR (Center for International Financial Analysis & Research, Inc.) has attempted to control for differences across countries in accounting standards.

Volume 2 contains the detailed information for each of the firms included in the resource. This information has been extracted primarily from annual reports and secondarily from various governmental filings. Annual financial information for as many as five consecutive years is presented for each firm (1990 is the most recent year). The annual data encompass 12 key financial ratios which are characteristic measures of profitability, liquidity, and risk. Four financial measures, in U.S. dollars, are also presented for 1990 regarding revenue, net income, total assets, and market value. Six similar measures are presented for up to five years in the domestic currency of each firm. Other details presented for each firm include the mailing address, number of employees, principal products, major shareholders, and stock exchange listings. In Volume 2, listings are presented alphabetically, by country. However, an index also provides a list of all the covered firms by industry.

Most of the space in Volume 1 is dedicated to the presentation of company rankings on a variety of dimensions. Before ranking the firms, CIFAR selected an elite 500 of the 7,500 plus companies listed in Volume 2. These 500 firms were selected based on relatively high rankings in terms of profitability, size, and stability. The precise selection process is not revealed in the publication. These 500 companies are ranked according to various size measures such as net income and number of employees. Also, these firms are ranked by size measures within industries and countries. In total, the companies are ranked in 15 different ways. Beyond the individual company rankings, average financial ratios are presented for all 7,500 companies by industry and by country. These ratios are presented by year for the period 1986–1990.

Two appendices appear in Volume 1. The first lists major differences in accounting treatments across the 48 countries for determining asset and expense amounts. Among the areas addressed are: inventory accounting, goodwill treatment, long-term equity investments, and revaluation of fixed assets. The second appendix contains the names and addresses of major stock markets for 44 countries as well as other summary information on capital markets in each country.

While this publication is no match for the electronic global data bases for research involving large samples and utilizing many financial variables, it has a role in accounting research. For example, this publication may be useful in locating addresses for mail surveys. Also, because the listings are arranged by both country and industry, it may be utilized to identify sample firms for international studies. Finally, in situations where hand coding of data is not an ominous task, limited data for individual firms is available in this resource and can be extracted with moderate effort. This option would be particularly appealing in situations where cost or other considerations limit the researcher's access to electronic databases. In this use, the fact that certain data appear in U.S. dollars for all firms may be a significant benefit. The researcher may be able to avoid the arduous tasks associated with currency conversion rates.

MICHAEL R. KINNEY
Assistant Professor of Accounting
Texas A&M University

PETER BOYS and JOHN FREEAR, editors, *Accounting History 1976–1986: An Anthology* (New York: Garland Publishing, Inc., 1992, pp. xii, 402, $87.00).

The anthology consists of 26 articles which were published in the now defunct British journal, *Accounting History*. The journal, published intermittently at minimal cost over a ten-year period, had a limited circulation since it was issued free to members of the England and Wales Accounting History Society. In publishing this anthology, the editors' objective is to have certain selected journal articles "reach a wider audience" (p. ix).

No central theme exists, but the articles are separated into eight different categories including general and miscellaneous. Other categories include historical methodology, auditing, firm/industry studies, corporate accounting, education, and biographies.

The articles that are most useful to accounting historians pertain to studies of accounting practices in various eighteenth and nineteenth century British industries. These studies include firms in industries such as iron, textile, brewing, baking (guilds), and railroad (audits). Two additional articles, one on accounting in bankruptcy proceedings, plus another on audit reports, provide the reader with an insight into complexities of accounting within an important time frame of the industrial revolution.

Further, an article researching the accounts of a medieval Venetian trader, one concerning eighteenth century British income tax policies, and another on Daniel Defoe, the accountant and author, also make for interesting reading. Articles on unique topics such as accounting for the American Revolutionary War, from the British perspective, and the earliest apprentices (1854–1863) serving Edinburgh's chartered accountants are especially insightful. The world's first professional accountants' organization received its royal charter in Edinburgh in 1854.

The overall quality of the 26 articles that make up the anthology is uneven. Some of the foremost authors in the field are included, but their contributions here are not likely to be viewed as among their best. The quality of the reproduction is dreadful. The original journals, according to the editors, were a mere collection of the authors' typewritten, double-spaced, manuscripts. And the book is a miniaturized copy of the original journals. It is the editors' stated wish that "the journal will be remembered for the quality of its articles rather than their presentation" (p. xi).

The editors are to be commended for their diligence in creating the original journal, as well as this anthology. The anthology contains many scholarly papers, but some that are of questionable merit. The authors' efforts probably will result in the papers reaching a wider audience of accounting historians.

HANS V. JOHNSON
Professor of Accounting
Arkansas Tech University

WILLIAM J. BRUNS, JR., editor, *Performance Measurement, Evaluation, and Incentives* (Boston: Harvard Business School Press, 1992, pp. vi, 304, $39.95).

Performance, Measurement, Evaluation, and Incentives is one of four books from a series in accounting and control at the Harvard Business School. The series editor, Robert S. Kaplan, and the board of advisers have selected some well-known international researchers.

This book is a collection of ten papers presented at a colloquium held at the Harvard Business School June 25–26, 1990. The colloquium focused on problems and issues in the measurement and evaluation of performance and the management of incentive compensation systems.

There are three sections: "Performance Evaluation and Incentives"; "Performance and Organizations"; and "Performance Measurement and Productivity." "Performance Evaluation and Incentives" contains six papers dealing directly with procedures and systems for evaluating performance in organizations. Four of the papers are case studies pertaining to development, evaluation, problems, and effectiveness of attempts to evaluate performance.

"Performance and Organizations" and "Performance Measurement and Productivity" each contain two papers. The lack of papers in these sections is a definite weakness in the discussion of these areas. The papers in "Performance and Organizations" discuss several key issues that should be discussed by organizations before implementing a performance evaluation and incentive program. The "Performance Measurement and Productivity" section raises questions regarding measurement performance in two different settings.

The editor implies in the introduction that the colloquium and the papers in the book were international in scope, because the manuscripts represent 16 professors from universities in the United States and Europe. This is deluding in that only two manuscripts are actually written by faculty from universities in Europe (Lancaster University and University of Siena). In fact, 60 percent of the papers are written by authors at only two different universities. In addition, although this research area overlaps labor economics, only one labor economist is represented in the book. Thus, the articles provide a rather narrow view of the research area and may not be international in scope.

Lastly, while the book does have some shortcomings, the editor is commended for successfully bringing together ten quality manuscripts in a new and exciting area of accounting research. While the papers do not answer all of the questions, they do provide some stimulating research issues. Persons in accounting or labor economics interested in performance evaluation and incentives, as well as performance measurement, will find this book useful.

MARVIN L. BOUILLON
Associate Professor of Accounting
Iowa State University

T. E. COOKE and M. KIKUYA, *Financial Reporting in Japan: Regulation, Practice and Environment* (Oxford: Blackwell Publishers Ltd., 1992, pp. xii, 356, $89.95).

Japanese financial reporting practices have long been a mystery to researchers, corporate managers, financiers, and investors. The striking contribution of this book is that the authors describe Japanese reporting practices within the Japanese environmental context rather than from the more popular U.S. or U.K. perspective. Another attractive feature of this work is that the authors have been able to successfully separate the discussion of accounting regulation from the discussion on accounting practice, providing the reader with a clearer picture of how the two intersect. Upon finishing this book the reader has a better, and fairly complete, understanding of the Japanese financial reporting environment as well as generally accepted accounting principles as regulated, adopted, and practiced in Japan. The reader also has new insight into the information disclosure issues of Japanese corporations and some grasp of the unique environmental factors that continue to influence accounting and reporting.

The book is logically divided into two parts. Part I provides a thorough discussion of the evolutionary environmental factors that have influenced the development of Japanese reporting practices, e.g., forms of business, users of the information and their needs, government regulation and the accounting regulatory environment, the accounting profession, various international influences, educational factors, culture, politics, and economics. Part II consists of an analysis of accounting principles and practices based on a survey of 48 Japanese corporate annual reports. In addition, this section of the book provides a wealth of information about currently practiced Japanese GAAP.

While the reader may find Part I a bit dry, this section is informative, educational, and necessary. It provides an important foundation for those readers interested in conducting research in this area. The chapter on stock exchanges and regulation is both relevant and timely, especially given the current interest in analyzing the efficiency of foreign stock exchanges and the publicity about Japanese P/E ratios. This section covers the organizational structure of the exchanges, while providing a discussion of the growth of the Japanese equities market. The authors offer a brief analysis of differences in P/E ratios with some explanation as to why these differences exist without attempting to offer answers to everything. They provide valuable historical information with ample charts and graphs for those preferring a visual walk through history.

The real strength of the book lies in Chapters 10 and 12, for if readers read only these two chapters they would have a solid Japanese financial reporting foundation to build upon. Chapter 10 covers disclosure principles and provides an overview of Japanese reporting as required by the Commercial Code and the Securities and Exchange Law. Chapter 12 reports the results of a descriptive survey of the content of 48 Japanese corporate annual reports; both listed and unlisted corporations are included. Information reported for the Comerical Code vs. The Securities and Exchange Law vs. the English language version is compared and some very revealing results are reported in the conclusions.

A tremendous amount of invaluable information is provided in 100 pages of appendices. There the reader finds an extensive comparison of reporting practices in Japan, the United Kingdom, and the United States. A chronology of accounting in Japan will appeal to those interested in a historical perspective. The book is an excellent reference work for the novice or the accomplished international researcher and is not suitable to be used as a text. There is a reference and further reading list at the end of the book that appears to be extensive and exhaustive. No bibliography is provided at the end of each chapter which would have been useful to the researcher.

In summary, the book has technical and research content and is a thorough, timely, and valuable contribution to the literature on Japanese financial reporting practices.

HELEN GERNON
Professor of Accounting

University of Oregon

GRAEME W. DEAN, MARC P. JOYE and PAUL J. BLAYNEY, *Strategic Management Accounting Survey: Overhead Cost Allocation & Performance Evaluation Practices of Australian Manufacturers* (Sydney: University of Sydney, 1991, pp. x, 144, A$25.00, paper).

This monograph presents a descriptive data base of overhead cost allocation and performance evaluation practices of Australian manufacturers. The survey is an extension of an earlier study by Joye and Blayney (see review in the January 1992 issue). While the earlier study covers a variety of management accounting issues, the current study focuses on overhead allocation and performance evaluation practices. As such, the subtitle of the book aptly captures the essence of the contents—but not the main title. The bulk of the book is devoted to data summary and cross tabulation. The reader will find relatively little textual material.

Out of over 2,000 manufacturers sampled, 313 provided usable responses. For readers interested in comparing practices in the U.S. and Australia, the vast majority of the firms sampled would be considered small and medium-sized by American standards. Of the 31 items on the questionnaire, 14 are related to the organizational and market characteristics of the sample firms, such as size, industry, diversification, ownership, etc., which serve as the categorical variables for cross tabulating responses to each of the remaining 17 questions.

In overhead cost allocation, the major findings include: (1) that allocated cost-based information is used almost exclusively for business decisions; (2) that prices are a function of costs or other factors, irrespective of whether managers perceive their firms to be price takers or price makers; (3) that overhead costs are allocated because the production "divisions would incur such costs if they were independent units" and "to make divisional managers aware that central services costs exist"; and (4) that the majority of respondents are satisfied with their current overhead allocation methods.

In performance evaluation, the major findings include: (1) that *accounting* rather than *non-accounting* data are predominantly used for performance evalution; (2) that accounting data are more likely to be used to evaluate upper-level managers and non-accounting measures are more likely to be used at lower levels; (3) that non-accounting indicators were found to provide *more timely information* than accounting-based measures; (4) that it is more common for Australian managers to use *absolute* (compared to predetermined objectives) rather than *relative measures* (compared to other managers) of performance; and (5) that over ⅓ of the upper-level managers were *dissatisfied* with their performance-evaluation system.

The majority of the book is devoted to an exhaustive cross tabulation of every item against the 14 categorical variables. This section is not particularly useful, as there is no statistical test to examine if any of these categorical variables contributed to any differences in responses. If there is any statistical difference, it should be brought to the attention of the reader; if not, there is no need to show extensive cross-tab tables.

The most useful section is probably the summary score for each questionnaire item. It allows readers to draw their own conclusions as well as seeing the anomalies in responses, such as including nonmanufacturing costs in inventory valuation for financial reporting purposes (over 20 percent of respondents) and checking either *every day* or *every week* for the *frequency* of measurement of *all inclusive income, operating income, return on investment, customer satisfaction,* and *market share.* The data in this section also led this reviewer to question the authors' conclusion that "the evidence provides limited support for criticism that far too much emphasis is placed on the short term by managers" (p. 28). The response summary shows that most of the short term measures are *operational measures* which are intended to provide "more timely information" (p. 5) and therefore are expected to be for a shorter time period.

In summary, with data anomalies in mind, the reader may find useful information about Australian manufacturers' overhead cost allocation and performance evaluation practices if one is prepared to draw his/her own conclusions.

SHU S. LIAO
Professor of Accounting

Naval Postgraduate School

ADOLF J. H. ENTHOVEN, JAROSLAV V. SOKOLOV, and ALEXANDER M. PETRACHKOV, *Doing Business in Russia and the Other Former Soviet Republics: Accounting and Joint Venture Issues* (Montvale, NJ: Institute of Management Accountants, 1992, pp. xv, 259, $29.95, paper).

This book is the culmination of a research study conducted for the Institute of Management Accountants. The principal aim of the book is to describe and provide insight into the system of accounting and auditing in the former Soviet Union. Although the study was commissioned before the breakup of the Soviet Union, its results continue to be of current interest as the Soviet rules have been adopted by the Russian Republic and are likely to be adopted by other members of the Commonwealth of Independent States.

The book is organized into three parts. Part I describes the Soviet accounting system. A chart of accounts is the key element of Soviet accounting with the balance sheet and other reports flowing from the accounts. In Chapters 1 and 2, the current Soviet chart of accounts (which is reproduced) and the methodology of financial accounting are described, and illustrative financial statements are provided. Perhaps of greatest value is the comparison of Soviet and Western accounting. Here the reader discovers, for example, that the principle of conservatism is not followed in Soviet accounting. Soviet cost accounting is described in Chapter 3 by means of a *T*-account illustration.

In Part II, the authors outline the process of negotiating and establishing a Soviet joint venture (Chapter 4). This provides a valuable introduction for anyone interested in establishing such an operation in the former Soviet Union. This section also includes a concise and easy-to-follow description of joint venture taxation complete with a flowchart and a tax computation format. Given the unfamiliar nature of some aspects of the Soviet tax system, this discussion would have benefited from the use of an illustrative example. In discussing joint venture accounting in Chapter 5, emphasis is placed on the preparation and format of the balance sheet. There is no direct discussion of joint venture income measurement, but the format of the profit and loss statement is provided. The authors indicate that a substantial restructuring of the financial statements is necessary to answer such questions as "what is the current ratio." A guide to such a restructuring would have been very useful to really understand the differences that exist between Soviet joint venture and Western accounting.

In Part III, the authors provide their predictions and suggestions for how Soviet accounting must evolve to realize its full potential in a market-oriented economy. Not surprisingly, the general prescription is that accounting must change from an instrument used to exercise control over economic activities and become a tool useful in management decision-making. It is suggested that a new generation of accountants must be created for this evolution to take place.

To help the reader better understand and appreciate the Soviet system of accounting, a glossary of Soviet accounting terms is included after Part III. In addition, 58 pages of appendices detailing (1) socio-political aspects of change in the Soviet Union, (2) the historical development of Soviet accounting and auditing, and (3) accounting education in the Soviet Union are also provided.

This book is a valuable handbook for anyone interested in knowing more about Soviet accounting, including academicians, practicing accountants, and persons interested in doing business in the former Soviet Union. It provides current information on financial accounting and reporting, cost accounting, auditing, taxation, accounting education, and joint venture formation in the former Soviet Union. For accounting faculty, it provides a useful reference for incorporating information on Soviet accounting into courses dealing with comparative accounting systems.

TIMOTHY S. DOUPNIK
Professor

University of South Carolina

MARTIN N. HOOGENDOORN and LILLIANE B. BIJL, editors, *Financial Reporting in the Nineties: Regulation and Innovation at a Crossroads* (Amsterdam, The Netherlands: Foundation for Annual Reporting, 1992, pp. 80, 47.50g, paper in English).

In October, 1991, The Foundation for Annual Reporting organized a conference on the future of financial reporting in The Netherlands. The title of this conference was "Financial Reporting in the Nineties—The Crossroad Between Regulation and Innovation." The purpose was to evaluate the present situation in the area of financial reporting in The Netherlands and to look especially at possible future tasks and at the future position of the Council for Annual Reporting (CAR).

The conference, chaired by Professor Traas (Amsterdam), included addresses by—inter-alia—Stephen Zeff (Rice University) and Arthur Wyatt, in his capacity as president of the International Accounting Standards Committee (IASC).

Some of the questions which the conference aimed to address were:

- Should regional and national standards be replaced by international standards?
- Should national standard setters wait for new requirements from regional or international bodies such as the European Community (EC), the IASC, or the OECD?
- What are the needs of users and regulators?
- How can annual reporting adapt to changes in the economic environment?
- How flexible are the European directives and national law?

Since the legal system of annual reporting is now fairly well complete in The Netherlands, the question became where to go from here; that is, should more detailed regulations be developed that bring more uniformity and a better guarantee for improvements in the quality of financial reports or should other roads be traveled, based upon the above questions raised.

The various discussion papers can be grouped under three headings:

1. developing accounting guidelines;
2. international comparability; and
3. compliance with accounting guidelines.

Regarding Item 1, involving *developing guidelines*, the Chairman set forth two alternatives: (1) the present legal system already specifies the minimum requirements, and (2) more detailed regulations that bring more uniformity. One speaker felt that detailed regulations have the psychological effect of diminishing the tendency to adopt new policies and therefore have an adverse influence on innovation. He favored a certain freedom, and in developing guidelines, the international choices on the level of detail will become more important. Steve Zeff mentioned that the expectation of unanimity between the three CAR constituencies in deciding accounting guidelines would be found unbelievable by bodies issuing standards in other countries; it meant an awful lot of compromising. As a result, people may have difficulty understanding the reasoning behind the statements.

In regard to *international comparability*, according to Zeff, The Netherlands was more serious about issuing guidelines as a consequence of the development of the IASC program in the 1970s. He considers the international influence to be healthy, without putting Dutch accounting in a straitjacket. Other speakers indicated that the EC Fourth Directive has been the direct cause of a major accounting reform in most Member States. The accounting directives are aimed at comparability and equivalence of accounting information and no attempt has been made to provide for uniform accounting rules. Considering the many differences of a legal, socio-economic and cultural nature between the EC Member States, it would have been very difficult to merge the different approaches into a single solution. Art Wyatt gave an overview of the activities of the IASC. He stressed the importance of international standards. Its main purpose is to reduce the uncertainties and inefficiencies in the capital market, thereby lowering capital and borrowing costs. Wyatt does not see the IASC and the EC as competitors; they both have the same end objective, although the method that is employed may differ. The work of the IASC benefits enormously from the work that has been done in the EC. However, he believes that in creating harmonization, one single set of standards is in the long run a better and more efficient alternative than mutual recognition.

In connection with *compliance*, several speakers stressed that historically the level of financial reporting in The Netherlands was high. However, Rinnooy Kan pointed out that some publications on annual reporting have expressed concern pertaining to compliance with law and accounting guidelines. However, he considered it doubtful whether a collective interest in quality control of annual reporting exists.

Some conclusions that can be reached from the conference discussions were:

1. The current condition of unanimity in accepting the CAR Accounting guidelines could hinder the standard setting process.
2. A certain backlog exists in CAR outputs, and more international cooperation using IASC standards as an input to the process may be a realistic way of decreasing the backlog.

3. International accounting standards setting is a very important ingredient to current national accounting standard setting.
4. The EC and the IASC have different ways of realizing international comparability: the EC emphasizes the acceptability of different alternatives with a common understanding, while the IASC is more conceptual and emphasizes the reduction of acceptable alternatives. Both the EC and IASC address the same end objective.
5. One of the dangers affecting the quality of financial reporting is the reduced influence of independent auditors owing to the commercialization of the auditing profession.

A critical question addressed was whether changes are to be expected in the near future in accounting standard setting in The Netherlands. Contrary to other countries, historically there has been no widespread criticism of financial reporting, and there were no financial reporting scandals that created the need for more, or modifications to regulation. No changes are expected in accounting standard setting in the near future. The CAR is bound to continue the codification of generally acceptable accounting practices.

The conference papers throw an interesting light on both the structure of financial reporting in The Netherlands, its current process and the future outlook for standard setting. This monograph is of considerable value for those interested in financial reporting outside our North American borders and those covering international accounting in academe or in practice.

ADOLF J. H. ENTHOVEN
Professor of Accounting
The University of Texas at Dallas

LEONARD LORENSON, *Accounting for Liabilities* (New York: American Institute of Certified Public Accountants, 1992, pp. xv, 173, $30.00, paper).

The Foreword of this book states that the book is intended "... to stimulate study and discussion of significant issues in financial accounting" (p. vii). The author has achieved this by presenting an extensive discussion and analysis of both the nature of, and reporting issues relating to liabilities. If readers are expecting a procedural handbook on different types of liabilities, they will be disappointed. This book attempts to identify and explain basic concepts and theoretical aspects of liabilities.

The conclusions of the study are that current practice (GAAP) for reporting of liabilities "... violate FASB's requirements that all such amounts faithfully represent attributes of assets and liabilities" (p. xiii). Lorenson reaches these conclusions by a careful, fundamental, and reasoned approach.

Using FASB Concepts Statements, Lorenson concludes that liabilities should be stated at their monetary attribute and that attribute should relate to events external to financial accounting. On this basis, Lorenson rejects current GAAP which requires that a loan be reported at the present value of the future payments. In this rejection, he is disagreeing with the proposed FASB "Present-Value-Based Measurements in Accounting" which has expressed the view that liabilities should be stated at present value, to the extent possible, and he is disregarding FASB Statement of Financial Concepts No. 5. Paragraph 67, which includes present value among the possible attributes which could be used for reporting liabilities.

The majority of the monograph is a discussion of what attribute should form the basis of reporting for each type of liability

Present value is discarded by the author as an appropriate attribute on the basis of his assumption, "A liability should be recorded at an amount of money that is an attribute of the liability outside of financial reporting" (p. 11). In eliminating this attribute Lorensen argues that, at dates subsequent to the incurrence of the liability, present value is not an attribute because it has no real-world referent.

Lorenson lists five possible alternative attributes at which liabilities could be reported: loan proceeds, the probable amount of the loan, the hypothetical proceeds of the loan, the creditor's acceptable early-discharge amount, and the risk-free funded amount. By a process of elimination, he concludes that virtually all liabilities should be reported at the risk-free funded amount unless the creditor's acceptable early-discharge amount is known and is a lower amount. The risk-free funded amount is the amount that would have to be invested in risk-free securities to make the payments.

I question why the risk-free funded amount is acceptable as an attribute when present value, which is obtained in a similar manner, is not. This infers that the funded amount has an external referent while present value has not. I cannot agree with that inference since inherent in choosing the funded amount is having to define the risk-free rate, a procedure which is controversial and may have no real-world referent.

There is no quibble with the elimination of loan proceeds and probable amount as possible attributes. These amounts are not reasonable amounts at which to state the liabilities. Hypothetical proceeds, which are analogous to reporting an asset at current replacement price, are eliminated as the author concludes that this is the same as reporting the liability at the value of a hypothetical loan, which is "... a liability it does not owe" (p. 46).

The author, concerned with choosing the appropriate monetary attribute for reporting purposes, uses fixed-payment loan liabilities as an example. The conclusions made at this point are then used for other liabilities such as credit purchases, leases, pensions, and income taxes. Controversial issues such as the reporting of stock appreciation rights and options and off balance sheet financing are not discussed.

In order to define a liability it is necessary to determine when an obligation is incurred. Addressing this issue, Lorenson takes a controversial stand stating that "... an obligation is incurred when performance becomes unconditionally required" (p. 5). This definition means that current GAAP for many liabilities is rejected. For example, Lorenson states that recognizing a liability for post-retirement benefits under FASB No. 106 results in a violation of the definition of a liability because the liability is recognized before it is incurred. Other examples, showing rejection of GAAP, are given.

Lorenson states that in determining GAAP, FASB has used judgment without previously developing a principle. He maintains that the obligation, and therefore the liability, is incurred when the last event occurs, which means that performance becomes unconditionally required. However, having developed that principle, the author immediately states four cases which are exceptions. This implies that, in this case, statement and adherence to a principle which has general application is difficult.

Although some may consider this a minor point, there is no index to this book. This is irritating and frustrating for the reader. The very nature of the topic is such that the reader may wish to use this book as a reference text where inquiry would be by subject. The lack of an index makes this difficult.

For those interested in basic concepts, fundamental reasoning, and the use of those in authoritative pronouncements, this monograph presents an alternative analysis with provocative conclusions. Although, in parts, this book is difficult to follow, it makes interesting reading as the author's views diverge from the views of most, if not all accountants. However, in my opinion, before Lorenson's arguments can be accepted by the profession, there has to be further discussion of the issues and the alternatives, and possibly, empirical evidence, since the resolution of such difficult issues is not likely to be resolved via *a priori* reasoning alone.

EILEEN PEACOCK
Assistant Professor of Accounting
Oakland University

SHARON M. MCKINNON and WILLIAM J. BRUNS, JR., *The Information MOSAIC* (Boston: Harvard Business School Press, 1992, pp. viii, 259, $24.95)

The Information MOSAIC contributes to the continuing indictment of managerial accounting by reporting the results of an exploratory survey conducted to determine the utility of accounting information to operating managers in manufacturing firms. In addition to reporting survey results, the authors make suggestions for increasing the relevance of accounting information. The reader is convinced that accounting reports are a minor subset of information used by managers, but is uncertain as to how reports can be made more meaningful.

The authors conducted a field study consisting of interviews with production, marketing, and other managers in twelve manufacturing corporations located in the United States and

Canada. The interview questionnaire focused on the information used by managers and the source of the information used in order to assess the relevance of management accounting reports to routine managerial decision making. The impact of management background and training and data communication technology on the use of information was also explored. The contribution of the researchers is the identification of information considered essential by managers and the manner in which they obtain and use this information.

The early chapters report on the specific information needed and used for production, sales, and marketing decisions. The impact of the characteristics of the production process, the size of finished goods inventory, and customer relationships on information needs is detailed. The discussion is thorough, easy to read, and supports the conclusion that managers use a mosaic of information most of which is not provided by the accounting system. Physical quantities or counts were regularly identified by respondents as critical factors in operating decisions. These measures are provided in a timely manner by non-accounting personnel and focus on production and marketing activity. Information meetings, plant tours, telephone conversations, and customer contacts provide the bulk of the information needed and used. The findings of other researchers on related issues are effectively incorporated into the discussion. An extensive list of references is provided, along with appendices describing the methodology of the study and the field questionnaire employed by the authors.

In the final two chapters the authors discuss the implications of their findings for management accounting. Accounting data is characterized as lacking timeliness. "A warehouse so full that it can receive no more finished product, demands a manager's attention much more quickly than a report on resources invested in inventory" (p. 8). In addition, the format of accounting information limits its usefulness. "Decisions and actions are based on physical quantities, not financial attributes" (p. 198). Doubt is expressed that accounting will be the sole or primary information system used by managers. Nonetheless, the authors reject the conclusion that additional investment in accounting systems fails the cost-benefit test. They contend that to permit the individual manager to determine his or her information needs and select the appropriate source of that information is to perpetuate the situation they found in practice. Rather, accounting information is central to evaluating the performance of managers over extended time periods. Suggestions for improvement to the management accounting process include development of a real-time data base, which the manager is able to access to meet specific information needs. The underutilization of electronic technology identified during the study is cited as another indication of the failure of management accounting to support operating management. According to the authors, management accounting must become more of a service function because of the diverse needs of managers who currently rely on personal information systems. After considering these recommendations, many of which have been made before, the reader is left to determine the manner in which the desired improvements to management accounting can be achieved.

In summary, the authors accomplished their objective of identifying the information needed by operating managers and the manner in which they obtain it. Accordingly, the results of this study will be of interest to both managers and managerial accountants.

FRANK R. PROBST
Professor of Accounting

Marquette University

YASUHIRO MONDEN, *Cost Management in the New Manufacturing Age: Innovation in the Japanese Automotive Industry* (Cambridge, MA: Productivity Press, 1992, pp. xxii, 176, $45.00).

Cost Management in the New Manufacturing Age: Innovation in the Japanese Automotive Industry is based on the author's study of leading Japanese automobile manufacturers. Combining the results of surveys and interviews, Monden has developed a concise description of the effects of innovative thinking on cost accounting, cost control and cost planning in Japan's automobile industry and augmented the text with case studies. Arranged in chapters, the topics flow as a product would flow through a three stage production line. Each stage includes information flow reinforced by a case study.

Stage one includes chapters two through five. Chapter 2 discusses the Toyota cost planning system which integrates cost management and product design and Chapter 3 concentrates on the Toyota cost management system. A framework of the Just-in-Time (JIT) production system developed by Toyota (based on Taylor's scientific management system and Ford's mass-assembly line) is presented in Chapter 4. Chapter 5 explores the relationship between JIT and cost accounting through a case study of the Daihatsu Motor Company's use of the Toyota production system.

Stage two, chapters six through eight, concerns profit/cost management systems. Chapter 6 concentrates on MRP, seiban and kanban production control methods. A case study of Kubota Company Ltd.'s profit and cost management system as implemented in its tractor production plant at Tuskuba is followed by a discussion on integrating a profit/cost management system into a responsibility accounting system. The effect of current changes in the production and sales environment on standard cost accounting is also explored.

Chapter 9 of stage three explains the full cost-based transfer pricing, risk-sharing, and risk-spreading behavior between suppliers and producers of automobiles in Japan. The final case, Chapter 10, examines the effect of a factory automation decision of the Daihatsu Motor Company as it reached for economic efficiency in its operations.

Anyone interested in a basic understanding of the cost accounting, cost control, and cost planning methods used by the Japanese automobile industry to achieve global recognition will find this book of interest. The glossary of basic terminology used in discussing the new manufacturing environment is excellent. Since five chapters are based on previously published work by the author, those familar with the literature would probably find the cases of most interest.

From an academic perspective, *Cost Management in the New Manufacturing Age* would make an excellent supplemental text for a Managerial Accounting course on both the undergraduate and graduate level. Unfortunately, it is not in paperback.

HELEN M. TRAUGH
Associate Professor of Accounting
University of Alabama at Birmingham

ASHEQ RAZAUR RAHMAN, *The Australian Accounting Standards Review Board: The Establishment of its Participative Review Process* (New York: Garland Publishing, Inc., 1992, pp. xxiii, 540, $73.00).

This book reports a descriptive study of the development of the Australian Accounting Standards Review Board (ASRB). The author uses a public choice model as his framework for explaining the reasons for the ASRB's creation, the process of its creation, and the choice of its unique set of functions. In particular, the author develops his model from the "Theories of Economic Regulation" developed by Stigler (1971), Posner (1974), and Peltzman (1976). The book also explains the functions of the ASRB, but it does not explain the standard setting process of the ASRB.

The author was inspired by the work of R. Walker (1987) who used the "regulatory capture" argument to explain the accounting profession's use of influence during the creation of the ASRB for the purpose of controlling the ASRB. The Walker piece did not examine the involvement of interested parties other than accountants in the formation of the ASRB; the current work does examine the involvement of other groups in the development of the ASRB. Using the public choice model, the author identifies the "demanders" for the ASRB as bureaucrats, the accounting profession, and companies. The identified "suppliers" of the ASRB are the politicians who served on the Ministerial Council for Companies and Securities. The author's hypothesis for examination in this study is: "In order to ensure that they played an influential role in the function of the ASRB all identifiable interest groups actively participated in the Board's establishment" (p. 5).

At its beginning, the book describes the organization and functions of the ASRB and makes the point that while the ASRB was organized as an independent board, it was the outcome of a multiparty process and is subject to substantial influence from several interest groups. A history of the development of Australian accounting regulatory policies is presented to explain why accounting standard setting bodies gained prominence as a source of disclosure regulation.

Competing theories relevant to regulation in accounting—market theories, capture theory, and public choice theory—are examined and a model to examine the creation of the ASRB is developed. Publicly available data (e.g., government documents and newspaper articles) are analyzed to explain the environment which led to government backed mandatory accounting regulation. Publicly available data are analyzed to explain the behavior of the demanders and suppliers in the process of creating the ASRB and to explain the establishment of the ASRB's functions.

The book is the author's dissertation. It is well organized but not well written and does not appear to have had the benefit of careful editing by the publisher. Such editing could have helped greatly to improve the clarity of presentation by avoiding pronouns without antecedents and nuisances such as a repeating line, a column of numbers in a table which does not foot, and so on. Those most likely to find this book useful include researchers interested in public choice theory; historians interested in the developmental history of the ASRB; those who would like to compare the development of American standard setting boards with experiences in other countries (perhaps to ascertain the extent of inclusion of all interested parties in the development of a standard-setting board and their participation in decision making processes); and those interested in a description of the functions of the ASRB.

W. T. WREGE
Associate Professor

Ball State University

ROBERT ROSENTHAL, *Meta-Analytic Procedures for Social Research, Revised Edition* (Newbury Park: Sage Publications, 1991, pp. x, 155, $13.95).

In recent years, meta-analysis has become popular as a method for generalizing from the findings of a cross-section of social and behavioral studies. The purpose of this book is to help "readers to conduct their own meta-analyses and to evaluate more wisely the meta-analyses conducted by others."

This edition revises the first edition by including some newer procedures, such as an effect size indicator for one-sample data, a new coefficient of robustness, and estimates of the file drawer problem. Although this revision is rather minor, the book rigorously describes meta-analytic techniques. The first chapter introduces the basic issues involved in summarizing a set of studies; it describes early applications, current status, and empirical evaluations of meta-analytic procedures. Chapters 2 and 3 introduce the basic concepts behind meta-analytic procedures. Defining a study's "results" as the relationship between variables, the relationship between tests of significance and estimates of effect size are demonstrated. Various types of effect size estimates and adjustments of these estimates are also included here. Retrieving and assessing research results are also discussed. Chapters 4, 5, and 6 provide the details for various quantitative techniques for meta-analytic procedures. Using instructive illustrations, the book demonstrates how to compare and combine the tests of significance and the effect size estimates from multiple studies. An effect size indicator for one-sample data is also introduced in Chapter 6. The final chapter discusses some criticisms of meta-analytic work.

The book has accomplished its intended purpose. It is well organized, and adequately describes the various procedures with illustrations. In sum, the book is valuable for individuals interested in the application of meta-analytic procedures.

SOON-YONG KWON
Assistant Professor

Yeungnam University

M. MAXWELL SPENCER, *The How & Why of a Tax on Wealth* (Saratoga, CA: R & E Publishers, 1992, pp. xxxi, 172, $14.45, paper).

The author's stated purpose is to present the case for a tax system that substitutes a tax on wealth coupled with a flat tax on non-saved income for current U.S. taxes including business, payroll, sales, property, transfer, and other taxes. The author's proposed wealth tax is equal to the real interest rate plus one percent multiplied times an individual's net worth. The author states, "In more Marxian language, we are proposing that we appropriate for public benefit the pure interest yield on capital, plus a little more" (p. iii).

In book one, the author lays down what he calls basic principles. Included in this section are discussions of the basic concepts such as the difference between the meaning of interest and *real* interest and the difference between tax avoidance and tax evasion. Although the author frequently refers to Marxism, he points out that support for the philosophy has declined primarily because, over the last fifty years, people have had the opportunity to empirically compare the two diametrically opposite systems. Nevertheless, the author says, "[W]hile private business is indisputably good from the standpoints of innovation, productivity, and rising living standards for the workers, its reappearance in the Communist world raises again the age-old question: What do you do about the special privilege that accrues to private wealth, once it is accumulated and passed on from generation to generation?" (p. 33).

The author estimates that his proposed net worth tax will raise about $2 trillion per year of tax revenue reduced by some amount for tax evasion. He says, "We should not push too far with negative yield through net worth tax, not so much because it may affect people's willingness to save, but because it may affect people's *willingness to comply*. We do not want a tax so onerous that people say: regardless of penalties, we'll hide our wealth, and we won't report it" (p. 49). The author apparently assumes that his proposed tax will not have this undesired result. Under his proposed system, however, the average wealth holder will be one percent better off hiding wealth than by bringing the wealth into the open and paying a wealth tax that is higher than the average return he or she can expect on investments. Over time, wealthy individuals who comply with the proposed tax law will see their net worth erode as they pay a tax greater than the income they can, on average, earn by investing their wealth. His estimate of tax revenue under his proposed system is based on the capitalization of GNP. Much of GNP, however, is produced by labor, not capital. One must assume that an author who frequently refers to Marx does not plan to tax human capital. Therefore, his estimated $2 trillion of annual revenues under the system is probably several times the actual amount that would be generated.

Book two deals with implementation. The tone of this book is set in the opening sentence, "From every individual, we want a Statement of Net Worth at the year end, showing all assets and all liabilities" (p. 91). The author perhaps understates the magnitude of such a requirement by saying, "[D]ifficulties in personal statements arise today..." (p. 93).

The flat *income* tax proposed by the author is imposed on "nonsaved income" or in other words, a consumption tax. The author proposes a $1,000 per person credit against the tax. Although the author proposes a flat tax, no rate is recommended. The author argues that taxing consumption instead of income, will encourage savings. The author offers no real support for his position that the incentive to save associated with a consumption tax will more than offset the disincentive associated with the tax on wealth. The proposed tax will undoubtedly have an extraordinary impact on investment, incentive, consumption, prices, property values, and income. The author fails to offer any integrated theory that explains exactly how each of these will be affected by the proposed tax system. There is no comprehensive estimate of how much revenue will be produced by the proposed tax system.

The author concludes with these words, "A restructured system can give labor over time the expropriation sought by the revolutionaries. Yet such expropriation is entirely consistent with a dynamic capitalism, where loss and profit incentives make the privately operated means of production function efficiently and innovatively, with resultant rising standards of living for the working population" (p. 154).

The author points out problems with the current income tax such as its complexity, the inequity associated with noncompliance, or perceived problems (the nondeductibility of child care). He suggests that his tax is better because it will not have these weaknesses. No evidence is offered in support of the author's claims that his proposed system will be less complicated, more equitable, or free from perceived problems. Would a deduction for child care under his system be better than the current credit? If Congress drafts the new law why does he expect the result to be less complicated or more equitable? The author accepts as equitable a tax system that imposes on a retired couple an annual tax that exceeds their income from their life savings. One must wonder why he believes a system that involves filing both an annual income and a net worth return as less complicated than the current system. Anyone who has ever prepared an estate tax return is aware of the complexity of a net worth tax.

Readers may find the book interesting, but far from complete. It lacks persuasive logic, a consistent theory, supporting evidence, and a detailed explanation of the proposed tax. It appears to be an attempt to propose a "modern" alternative to "historical" Marxism. Perhaps the reason the author's proposal lacks so many essential elements is that it, like the predecessor Marxist philosophy, is a fundamentally deficient system. One has to wonder whether a tax system that imposes an average tax rate on income from capital of over one hundred percent can reasonably be called capitalistic.

DALE BANDY
Professor of Accounting
University of Central Florida

TANG YUN WEI, LYNNE CHOW, and BARRY J. COOPER, *Accounting and Finance in China: A Review of Current Practice* (Hong Kong: Longman Group [Far East] Ltd., 1992, pp. x, 224, $55.00).

China is a huge country where an estimated nine million people currently work in accounting-related occupations. The accounting profession in China is experiencing a period of development and change, as it strives to raise its standards. This book attempts to provide an insight into the role and impact of accountancy as a discipline in China. It succeeds in that attempt.

The authors provide a comprehensive description of the accounting environment in China, showing it as a social science, developing within the socio-economic framework of the country. The enterprise reform and open door policies implemented in China over the past ten years have encouraged more freedom in the making of business decisions. The major issue, as yet unresolved, is how to account for the proprietary rights over state assets; here the government in its three different capacities as administrator of society, economic controller of the nation, and representative of state assets looks to the accounting profession to lead in this area.

In their early chapters, the authors succeed in portraying these challenges as the proper remit of the accounting profession. In particular, their discussion of the concept of "fund," "fund source," and "fund application" is admirable, and a model for policy makers elsewhere in the world who have to cope with the contents of statements of changes in financial position. The difficulties faced by accountants in accounting for state industrial and state commercial enterprises, and especially for collective enterprises, will be most illuminating for those from a capitalistic, for-profit accounting background.

Against this opening tableau, the section dealing with contemporary accounting issues in China clearly represents a formidable challenge. The share capital system as exemplified by the activities of the Shenzhen and Shanghai stock exchanges, and the difficulty of accounting for changing prices within a concept of physical capital maintenance are dealt with in a way which provided this reviewer with a feeling of genuine awe. It is here where the meat of the book is to be found, and it makes for a worthy repast.

Sadly, but perhaps inevitably, the succeeding chapters appear thin and lightweight in comparison. Consideration of accounting standards in their infancy, government accounting, budgets and control and the taxation system must be given to enable a comprehensive picture of accounting and finance in China to emerge, but these sections are heavily descriptive and rather turgid. However, for readers whose primary motivation is to discover the implications for them of the tax regime on enterprises with foreign investment and on foreign enterprises and foreign nationals, then such information is given.

The book closes with discussions of the development of public accounting qualifications and firms, and of accounting education, both of which are areas of concern to likely readers of this review. In the former area, the Chinese Institute of Certified Public Accountants (CICPA) was established as recently as November 1988, and this has proved to be a landmark event in the development of the accounting profession in China. It is also noteworthy that the Chartered Association of Certified Accountants in the United Kingdom has undertaken a major initiative in working with the China Association for International Exchange of Personnel to train Chinese accounting staff to acquire the ACCA professional qualification. In the area of accounting education at secondary and tertiary levels much remains to be done. A recent nation-wide survey highlighted weaknesses such as narrowness of the subject matter taught, over-specialization, outdated textbooks, and unsatisfactory teaching methods. Such problems are not unknown in Western economies as well.

This book is useful for practicing accountants who have, or are likely to have, professional dealings with enterprises in China or with its emerging accounting profession. Accountants currently in Hong Kong might find it of special interest. Also academics and students who are interested in the study of comparative accounting systems will find this book of enormous value.

ROBERT A. LYON
Senior Lecturer in Accounting
University of Dundee

BOOKS RECEIVED AUGUST 1, 1992 TO NOVEMBER 1, 1992

NICHOLAS G. APOSTOLOU and D. LARRY CRUMBLEY, *Handbook of Governmental Accounting & Finance* (New York: John Wiley & Sons, Inc., 1992, pp. xii, 910, $105.00).

MICHAEL BROMWICH, *Financial Reporting, Information and Capital Markets* (London: Pittman Publishing, 1992, pp. viii, 376, £21.99, paper).

PETER CHALOS, *Managing Costs in Today's Manufacturing Environment* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1992, pp. xii, 283, $36.60).

ADOLF J. H. ENTHOVEN, JAROSLAV V. SOKOLOV, and ALEXANDER M. PETRACHKOV, *Doing Business in Russia and the Other Former Soviet Republics: Accounting and Joint Venture Issues* (Montvale, NJ: Institute of Management Accountants, 1992, pp. xv, 259, $29.95, paper).

JACK R. FAY, *Accounting Certification, Educational, & Reciprocity Requirements: An International Guide* (Westport, CT: Quorum Books, 1992, pp. xii, 301, $65.00).

S. J. GASTON, *Managing and Controlling Small Computer Systems Including LANS* (Toronto: The Canadian Institute of Chartered Accountants, 1992, pp. xix, 280, C$29.50, paper).

MICHAEL GIBBINS, ALAN J. RICHARDSON, and JOHN WATERHOUSE, *The Management of Financial Disclosure: Theory and Perspectives* (Vancouver: The Canadian Certified General Accountants' Research Foundation, 1992, pp. xii, 97, C$16.75, paper).

KENNETH S. HACKEL and JOSHUA LIVNAT, *Cash Flow and Security Analysis* (Homewood, IL: Business One Irwin, 1992, pp. xi, 298, $60.00).

HERMAN H. HALE and ROBERT M. ATKISSON, *Controlling Vendor Transactions: A Guide for Auditors and Accountants* (New York: John Wiley & Sons, Inc., 1992 pp. xi, 149, $49.95).

MICHAEL HALPERIN and STEVEN J. BELL, *Research Guide to Corporate Acquisitions, Mergers, and Other Restructuring* (New York: Greenwood Press, 1992, pp. xv, 208, $49.95).

CHERYL R. LEHMAN, *Accounting's Changing Role in Social Conflict* (New York: Markus Wiener Publishing, Inc., 1992, pp. xi, 174, $34.95).

SEAN MOONEY, *Insuring Your Business* (New York: Insurance Information Institute Press, 1992, pp. vii, 209, $22.50, paper).

J. AUSTIN MURPHY, *Research Solutions to the Financial Problems of Depository Institutions* (Westport, CT: Quorum Books, 1992, pp. xii, 156, $47.95).

THOMAS R. PRINCE, *Financial Reporting and Cost Control for Health Care Entities* (Ann Arbor: AUPHA Press/Health Administration Press, 1992, pp. xxxii, 612, $65.00).

PRICE PRITCHETT, *The Ethics of Excellence* (Dallas: Pritchett Publishing Company, 1991, pp. xviii, 35, $5.95, paper).

DAVID L. RADOS, *Pushing the Numbers in Marketing: A Real-World Guide to Essential Financial Analysis* (Westport, CT: Quorum Books, 1992, pp. x, 185, $45.00).

AHMED RIAHI-BELKAOUI and ELLEN L. PAVLIK, *Accounting for Corporate Reputation* (Westport, CT: Quorum Books, 1992, pp. xiv, 253, $49.95).

AHMED RIAHI-BELKAOUI, *Morality in Accounting* (Westport, CT: Quorum Books, 1992, pp. ix, 212, $49.95).

KJELL ERIK RUDESTAM and RAE R. NEWTON, *Surviving Your Dissertation: A Comprehensive Guide to Content and Process* (Newbury Park, CA: Sage Publications, Inc., 1992, pp. xi, 221, $16.95, paper).

V. SAMBAMURTHY and ROBERT W. ZMUD, *Managing IT for Success: The Empowering Business Partnership* (Morristown, NJ: Financial Executives Research Foundation, 1992, pp. 97, $15.00, paper).

WASSIM N. SHAHIN, *Money Supply and Deficit Financing in Economic Development* (Westport, CT: Quorum Books, 1992, pp. xiv, 231, $55.00).

M. MAXWELL SPENCER, *The How & Why of a Tax on Wealth* (Saratoga, CA: R & E Publishers, 1992, pp. xxxi, 172, $14.45, paper).

MILTON F. USRY, editor, *Mary Ball Washington Forum Series in Accounting Education Volume VI* (Pensacola, FL: University of West Florida, 1992, pp. 57, Free, paper).

WANDA A. WALLACE and S. DUANE SMITH, *Accounting for Options: No Longer an Option* (Vancouver: The Canadian Certified General Accountants' Research Foundation, 1991, pp. x, 192, C$19.75, paper).

PAULINE WEETMAN and SIDNEY GRAY, editors, *SSAP 19: Accounting for Investment Properties* (Edinburgh: The Institute of Chartered Accountants of Scotland, 1992, pp. vi, 163, £15.00, paper).

KIM WOO-CHOONG, *Every Street is Paved with Gold: The Road to Real Success* (New York: William Morrow and Company, Inc, 1992, pp. 254, $20.00).

STEPHEN A. ZEFF, FRANS VAN DER WEL, and KEES CAMFFERMAN, *Company Financial Reporting: A Historical and Comparative Study of the Dutch Regulatory Process* (Amsterdam: North-Holland, 1992, pp. xv, 410, $85.00).

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal.... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of the *Chicago Manual of Style* (13th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's International Dictionary*.

FORMAT

1. All manuscripts should be typed on one side of 8½ × 11″ good quality paper and be double spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To assure anonymous review, authors should not identify themselves directly or indirectly in their papers. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title, and affiliation, any acknowledgements, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: *All* pages, including tables, appendices, and references, should be serially numbered. The first section of the paper should be untitled and unnumbered. Major sections may be numbered in roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights, and measures. For example: *three days; 3 kilometers; 30 years*. All other numbers are expressed numerically. Generally when using approximate terms spell out the number, for example, *approximately thirty years*.

Percentages and Decimal Fractions: In nontechnical copy use the word *percent* in the text; in technical copy the symbol % is used. (See the *Chicago Manual* for discussion of these usages.)

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: a *well-presented analysis; re-form*. See *Webster's* for correct usage.

Key Words: The synopsis is to be followed by four *key words* that will assist in indexing the paper.

SYNOPSIS

A synopsis of about 400 words should be presented on a separate page immediately preceding the text. The synopsis also serves as the "introduction." The synopsis should be nonmathematical and include a readable summary of the research question, method, and the significance of the findings and contribution. The title, but not the author's name or other identification designations, should appear on the synopsis page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an arabic number and a complete title indicating the exact contents of the table or figure.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix *a*, *b*, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says. . . ."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by the *Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Dates of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) should be listed in chronological order of publication. Two or more works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2d ed. Homewood, IL: Dow Jones-Irwin.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, United Kingdom: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–36.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, University of Texas, Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be double spaced and numbered consecutively throughout the manuscript with superscript Arabic numerals. Footnotes are placed at the end of the text.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or other publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan, or the like) should be submitted.
3. Four copies should be submitted together with a check in U.S. funds for $50.00 for members or $100.00 for nonmembers of the AAA made payable to the American Accounting Association. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the Editor for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting.

"An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional, research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

APRIL 1993 PLACEMENT ADS

The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The REVIEW *is published on the first day of January, April, July, and October. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233. ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding. There is no charge for this service to AAA members and associate members seeking employment. Others will be charged a fee of $150 per insertion. ADS SHOULD NOT EXCEED 150 WORDS.*

#201 UNIVERSITY OF MARYLAND European and Asian Divisions seek excellent teachers for undergraduate courses on U.S. military bases overseas. Minimum requirements include a master's in accounting, one year of recent college teaching experience, and U.S. citizenship. Renewable annual appointments begin August 1993. Benefits include transportation and military base privileges. Frequent travel and the cost of schooling make these positions difficult for those with children. Send resumé to: Dr. Ralph E. Millis, University of Maryland University College, College Park, MD 20742-1642. Affirmative Action/Equal Employment Opportunity.

#202 NATIONAL CHUNG HSIN UNIVERSITY invites applications for tenure-track positions beginning September 1993. Candidates must have a Ph.D. or be near completion. Preference is for experienced individuals in areas of auditing, financial, managerial, or accounting information system. Necessary qualifications include commitment to teaching excellence and quality research. Teaching is at the undergraduate and graduate levels with a minimum of 9 hours per semester. Competitive salaries with excellent benefits are available. The deadline for applications is July 1, 1993. Please send a letter of application, vita, list of three references, and research proposals to: Professor Lee, Chair, National Chung Hsin University, 67, Sec. 3, Ming-Shan E. Road, Taipei, Taiwan R.O.C., Fax: 02-5060730. Review of applications will begin immediately and continue until positions are filled. Equal Employment Opportunity/Affirmative Action Employer.

#203 APPLICATIONS FOR POSITIONS IN ACCOUNTING, at all professorial ranks, are invited from qualified persons. A Ph.D. is required. Demonstrated ability or potential for effective teaching and quality research are important considerations. The University of Pennsylvania is an Equal Opportunity/Affirmative Action Employer. Interested persons should send their resumé including educational background, professional experience, publications, personal data, and, if applicable, a dissertation proposal and abstract to: Professor Robert E. Verrecchia, Accounting Department, University of Pennsylvania, Suite 2400, Steinberg Hall-Dietrich Hall, Philadelphia, PA 19104-6365.

#204 ACCOUNTING: MICHIGAN TECHNOLOGICAL UNIVERSITY anticipates a Fall, 1993 position of tenure-track full professor. Requires Ph.D., research and publication record, professional certification. Send resumé to: William Chapel, School of Business and Engineering Administration, MTU, 1400 Townsend Drive, Houghton, Michigan 49931. Michigan Technological University is an Equal Opportunity Educational Institution/Equal Opportunity Employer.

#205 ANDREAS SCHOOL OF BUSINESS, BARRY UNIVERSITY is seeking applicants for a faculty position in accounting in September, 1993. Rank and salary open, based on qualifications and experience. Ph.D or DBA, completed or near completion; professional certification is desirable. A strong student orientation and commitment to teaching are required. Although we are primarily a teaching institution, some research is expected for continuing contract. Responsible for teaching basic and advanced courses in accounting at the Bachelor and Master's level. Interested candidates should send a resumé to: Dr. Lloyd D. Elgart, Dean, Andreas School of Business, Barry University, 11300 NE 2nd Avenue, Miami Shores, Florida 33161. Position subject to funding. Barry is an Equal Opportunity/Affirmative Action Employer.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 443–465

Auditors' Belief Revisions and Evidence Search: The Effect of Hypothesis Frame, Confirmation Bias, and Professional Skepticism

Jeffrey J. McMillan
Clemson University
Richard A. White
University of South Carolina

SYNOPSIS AND INTRODUCTION: The primary purpose of this study was to investigate how auditors' belief revisions and evidence search are influenced by the frame of the hypothesis being tested and by confirmation bias and professional skepticism (conservative bias). This study is distinguished from prior research in three main ways. First, the effects of confirmation bias and professional skepticism on the complementary audit functions of evidence evaluation and evidence search are examined jointly in the same experiment. Second, unlike many studies which have examined belief revisions of auditors solely under an error frame (Ashton and Ashton 1988, 1990; Tubbs et al. 1990), this study examines the judgments and behavior of auditors operating with both environmental (nonerror) and error-framed hypotheses. Third, the full effects of confirmation bias and professional skepticism were enhanced by having the subjects establish their own hypothesis frame and likelihood assessments, rather than respond to preset conditions.

Confirmation bias implies that auditors may seek to confirm their hypotheses and so may favor information that confirms rather than refutes their initial assessments. This approach could lead to premature closure on a belief or hypothesis. Professional skepticism implies that auditors focus

We appreciate the criticisms and suggestions of the anonymous reviewers, associate editor, and editor. We also are grateful for the comments on earlier drafts of this article from Michael Bamber, Bryan Church, James R. Davis, Bruce Hadlock, Paul Harrison, Arnold Schneider, Brad Tuttle, and for the guidance of dissertation committee members Thomas Cafferty, Gene Chewning, James B. Edwards, and Alan Winters. Special gratitude is expressed to James B. Edwards, who designed the computer program necessary to carry out the experimental task. An earlier version of this article was presented at the 1991 Annual Meeting of the American Accounting Association.

Submitted August 1991.
Accepted December 1992.

more on error-related evidence. An approach that is too conservative may lead to the performance of unnecessary audit procedures and thereby reduce audit efficiency. Environmental conditions refer to economic changes, changes in the industry and geographic area in which the company operates, or changes in company policies regarding investment, marketing, and financing strategies. Error conditions refer to intentional or unintentional misstatements in the financial statements.

A field experiment was conducted in which auditors reviewed preliminary audit information and then indicated whether they favored an environmental or an error-framed hypothesis as the most likely explanation of an observed fluctuation in financial statement ratios. A likelihood assessment was assigned to the favored hypothesis frame, and the auditors were then asked to seek information to test their initial hypothesis from a list of audit questions. After evaluating the audit cues, the auditors updated their hypothesis beliefs and continued their evidence search.

The results of this study indicate that auditors reacted differently to audit evidence, depending upon the frame of the hypothesis they favored and their belief extremity. Specifically, auditors who favored the error frame reacted more strongly to both confirming and disconfirming evidence than did those who favored the environmental frame. Furthermore, in conformance with the findings of Ashton and Ashton (1988, 1990) and contrary to Bamber et al. (1991), the auditors were more responsive to disconfirming evidence than to confirming evidence when belief revision was measured with an absolute scale. However, when the relative change in belief revision was measured with a proportional scale, the magnitude of response with confirming evidence was not significantly different from that with disconfirming evidence.

The auditors' continued evidence search was conditioned by a conservative bias irrespective of the hypothesis frame favored or belief extremity; that is, their search strategy emphasized the uncovering of potential material errors (Smith and Kida 1991). Because the conservative bias was stronger for auditors who favored the error frame, confirmation bias may partially account for this effect by enhancing the emphasis on error. Conversely, confirmation bias may have somewhat weakened the effect of conservative bias for those auditors who favored the environmental frame.

Key Words: *Belief revision, Evidence search, Hypothesis frame, Confirmation bias, Professional skepticism, Conservative bias.*

Data Availability: *Data used in this study are available from the first author upon request.*

AN experiment to examine the effects of hypothesis frame, confirmation bias, and professional skepticism consisted of two phases. In the first phase, reported in McMillan and White (1992), we established the manipulation of hypothesis frame and examined whether an auditor's *initial* hypothesis affected the initial evidence search and reflected a confirmatory or conservative search strategy. In the second phase, we continued to investigate the effects these factors had on the auditor's decision making following the evaluation of audit evidence. Specifically, we examined

(1) the updating of the auditor's beliefs after exposure to four evidence cues and (2) the auditor's selection of additional audit evidence to continue testing the updated beliefs. The details of the experimental method employed in the study and the findings from the second phase are presented in this article.

I. Formulation of Research Questions

Hypothesis Frame and Belief Revision

An auditor's hypothesis concerning the fairness of a client's financial statements can take one of two broad positions or frames: (1) that a material error exists in the financial statements or (2) that no material error exists (i.e., changes between last year's and the current year's financial statements are attributed to environmental factors). As previously noted, many recent studies had subjects test hypotheses that were framed with an emphasis on material error. However, in the first phase of our experiment, a significant majority of subjects favored an environmental frame after the review of preliminary evidence (McMillan and White 1992). The positioning of all subjects with the same initial hypothesis orientation limits the investigation of the full effect that confirmation bias or professional skepticism may have on subjects' belief revisions. Auditors who initially favor a hypothesis that has an environmental frame may react differently to audit evidence than auditors who favor an error frame.

Auditors are trained to be sensitive to evidence that reduces the risk of failing to detect material errors in the client's financial statements. The accounting literature often refers to this behavior as professional skepticism or conservative behavior (Smith and Kida 1991). If professional skepticism influences the belief revision of auditors, subjects who initially favor an environmentally framed hypothesis should react less to confirmatory evidence than subjects who initially favor an error-framed hypothesis. Alternatively, if confirmatory behavior significantly influences the audit judgment process, subjects with an environmentally framed hypothesis may react in the same manner as those with an error-framed hypothesis (i.e., both overweight confirming evidence and discount disconfirming evidence). It is possible that confirmation bias and professional skepticism may complement auditors' belief revisions, counteract each other, or the effect of one could dominate the other.

Because of the limited evidence on the effects of confirmation bias and professional skepticism in an environmentally framed setting, the research hypothesis regarding belief revision and framing effect is presented in the null form.

> H1: The magnitude of auditors' belief revisions after the evaluation of evidence will not be affected by their initial preference for an environmental or an error-framed hypothesis.

Evidence Direction and Belief Revision

Research in decision making indicates that individuals are inclined to evaluate evidence in a way that confirms their hypotheses and will have a greater response to confirming evidence compared to disconfirming evidence (Einhorn and Hogarth 1985; Gorman 1986; Lord et al. 1979). Recent studies in accounting have been mixed with respect to whether auditors react more to confirming or disconfirming evidence. In experiments involving the evaluation of audit evidence, Ashton and Ashton (1988) found that auditors reacted more (i.e., had greater belief revisions) to disconfirming evidence than to confirming evidence when their likelihood assessment was set at 50 percent. This result contradicts prior decision-making research in nonaccounting areas. When the initial anchors were set at relatively extreme levels (20 percent and 80 percent), the magnitude of belief revisions was significantly greater for confirming evidence at the low anchor and significantly greater for disconfirming evidence at the high anchor. Ashton and Ashton's (1988) results supported Einhorn and Hogarth's (1985) contention that end-of-scale effects may dominate behavioral traits.

In a follow-up study, Ashton and Ashton (1990) found that disconfirming evidence caused greater belief revisions than confirming evidence for auditors but had the same effect on business executives. Asare (1992) also noted that inconsistent evidence was more persuasive than consistent evidence in an experiment involving auditors' assessments of a going-concern decision. In contrast, studies by Bedard and Biggs (1991) and Church (1991) found that auditors exhibited confirmatory tendencies in their evaluation of evidence. In both of these studies, subjects were allowed to generate their own hypotheses. In the studies by Ashton and Ashton (1988, 1990), the subjects were given the initial likelihood of an error occurrence. Church (1990) argues that confirmatory behavior may be affected by whether the initial hypothesis is generated by the auditor or inherited (i.e., suggested by client personnel, superiors, or decision aids). Bamber et al. (1991) provided subjects with an initial hypothesis to test and told them to assume an initial likelihood estimate of 50 percent. Their results reflected a sensitivity toward confirming evidence that is consistent with Church's findings, even though the subjects were furnished a hypothesis and likelihood estimates in a manner similar to that in the studies by Ashton and Ashton (1988, 1990).

The conflicting results may be attributed in part to whether the hypothesis and the initial likelihood assessment are generated or furnished. The findings of Butler (1986) suggest that positioning auditors in a judgment task with a "given" initial anchor may not be totally effective because auditors tend to bring their own internal anchor to the audit task. If a given initial likelihood anchor is artificially low, the measurement of reactions to confirming evidence may be overstated and that of disconfirming evidence understated. Ashton and Ashton's (1988) findings indirectly support this contention as reaction to confirming evidence decreased when the anchor treatment increased, particularly when it increasd from 50 to 80 percent. For experiments designed to imply a two-state action space, a 50 percent anchor may be too low. This contention was supported in the first phase of our experiment (McMillan and White 1992).[1]

Because of the conflicting results in the accounting literature regarding the effect of evidence direction on the magnitude of auditors' belief revisions, the reseach hypothesis is stated in the null form.

H2: The magnitude of the belief revisions of auditors who evaluate confirming evidence will equal the magnitude of the belief revisions of auditors who evaluate disconfirming evidence.

Continued Evidence Search

As discussed in the introduction, we examined the evidence evaluation and search phases of an audit investigation in one overall experiment because it is possible that certain decision biases or situational factors that are significant in one stage of the audit judgment process may not influence decisions in other stages. Presently, a unifying model that addresses the relationships among belief revision, evidence search, and the final audit decision does not exist. This study contributes to that goal by examining whether two factors (hypothesis frame and evidence direction) influence the belief revision and evidence search phases of an audit in a consistent manner.

---

[1] The mean initial likelihood established by the subjects was 70 percent which is consistent with prior studies that allowed auditors to set their own initial anchors (Kaplan and Reckers 1989; Tubbs et al. 1990).

Hypothesis Frame and Evidence Search

Snyder and Swann (1978) proposed that decision makers generally employ one of three information search strategies: (1) a search for evidence that confirms the hypothesis under scrutiny, (2) a search for disconfirming evidence, or (3) a "balanced" search that invests equal amounts of effort to uncover both confirming and disconfirming facts. Klayman and Ha (1987) supported the conclusions of Snyder (1981) that people in general engage in confirmation-prone search strategies.

Kida (1984) examined whether auditors' initial evidence search in performing an audit-related task exhibited confirmation-prone tendencies. He manipulated hypothesis frame and found that auditors in a failed-company treatment group selected more information items connected with failure than with viability, which implies confirmatory behavior. However, the viable-company treatment group also selected more information items connected with failure, which implies disconfirmatory behavior. Smith and Kida (1991) contend that this proneness toward failure information is consistent with the professional and legal environment of auditing and labeled it a "conservative" search strategy. In the context of the current study, conservative behavior suggests that auditors would be inclined to employ a search strategy that is biased toward the uncovering of material errors.

Because of the limited research on the effect of hypothesis frame on auditors' continued evidence search (i.e., the subsequent search after the evaluation of previously collected audit evidence), the research hypothesis is presented in the null form.

> H3: When selecting additional audit evidence, auditors' confirmatory and/or conservative strategies will not be affected by whether they favored the environmental or error-framed hypothesis.

Prior Evidence and Evidence Search

The effect of confirming or disconfirming evidence on auditors' continued search for evidence has not been examined. Conceptually, the prior literature could support an argument for a positive or a negative relationship between the direction of the prior evidence evaluated (confirming vs. disconfirming) and the type of subsequent evidence selected (environmental vs. error).

For example, if auditors have a propensity toward confirmatory behavior, have examined evidence that confirmed their hypothesis, and have acted in a consistent manner throughout the investigation, they should continue to direct their evidence search in a confirmatory manner. This behavior would indicate a positive relationship between prior evidence and the subsequent evidence selected. If, however, auditors initially receive disconfirming evidence, the resulting downward revisions in belief could result in a shift to the alternative frame. Auditors who change their hypothesis frame and act in a consistent confirmatory manner should seek evidence that confirms the alternative hypothesis frame. Thus, in this case, a negative relationship between evidence direction and evidence search would occur even though the search process is consistent with confirmatory evidence. If, however, conservative traits dominate, the above relationships may not follow or may depend on the frame of the hypothesis favored. Accordingly, the following null hypothesis was tested.

H4: The type of evidence selected by auditors during the continued search for evidence will not be affected by the direction of the evidence previously evaluated.

II. Experimental Procedure

The subjects consisted of 50 staff, 50 seniors, and 66 partners and managers who held full-time auditing positions in five regional CPA firms (18 percent) and three Big-Six CPA firms (82 percent). The Big-Six subjects were taken from a minimum of two offices for each firm in six different cities. The sample group was composed of 104 males and 62 females. All of the subjects held undergraduate degrees, and 37 had completed graduate degrees.

A contact person in each office randomly distributed and collected a packet of printed materials detailing preliminary audit information for each subject, along with a computer diskette to each subject.[2] The experimental task was controlled by a computer program that was both iterative and interactive. The subjects were provided preliminary financial and nonfinancial information concerning a hypothetical client. The financial information consisted of two years of summary data related to the balance sheet and income statement, financial ratio profiles for the prior year (audited) and current year (unaudited), and industry and projected ratios for the current year. Two of the three ratio profiles showed a substantial change between the last year's and the current year's ratios.[3] Half of the subjects received ratio profiles that presented the industry and the projected ratios as substantially the same as the client's current-year ratios, while the other subjects received ratio profiles in which the client's current-year ratios differed from the industry and projected ratios. The nonfinancial information described the client's internal control environment. The subjects who received similar financial information received a client portrayal of strong internal control. Conversely, the subjects who received the dissimilar financial information received a weak internal control portrayal.[4]

After studying the preliminary audit information, the subjects were asked to select the hypothesis frame that "most likely" identified the source of the fluctuation in the current year's financial statement ratios. One hypothesis was framed to suggest that environmental factors were responsible, while the other hypothesis frame suggested that material errors in the unaudited statements were the cause of the ratio fluctuations.[5] Following this selection, subjects were asked to indicate the likelihood, from zero (no chance) to 100 percent (completely certain), that the hypothesis frame selected represented the true state. Following the formation of their initial beliefs, the subjects were asked to review a "menu" of 16 audit questions provided in the audit packet. Two types of audit questions were randomly positioned throughout the menu—eight environmental questions and eight error questions. The questions and their classification were derived by Kaplan and Reckers (1989) and Libby (1985), and all were reviewed by several CPAs and determined to be plausible explanations for the ratio changes.[6] From the menu of audit questions, the subjects rank-ordered the first four questions they would initially seek answers to in an effort to explain the marked change in the financial ratios.

---

[2] A letter from each office's audit partner asking subjects to complete the research project in a timely manner was included in each packet. A brief set of instructions informed the subjects how the materials were to be used and that the study examined strategies used by auditors and identified characteristics of audit decision making. Pretests were performed before the instruments were distributed.

[3] The financial ratio profiles consisted of three ratios—gross margin, current, and quick. The profiles were similar to the ones used by Libby (1985), Kaplan and Reckers (1989), and Libby and Frederick (1990). The marked change was between last year's and the current year's current and quick ratios.

[4] The preliminary audit evidence was provided so that the subjects would have information for an initial hypothesis. The purpose of the manipulation was not to test the particular effects of the financial and nonfinancial information; thus, a full factorial manipulation of the information was not germane to the study. The primary purpose of the manipulation was to generate a diversity of beliefs for two hypothesis frames (environmental vs. error) so that framing effects could be investigated. Comparative industry financial data, financial projections, and internal control environment are specifically identified by SAS No. 56 (AICPA 1988) as sources auditors should use when forming their initial expectations.

The above tasks concluded the first phase of the experiment. The second phase began with the subjects being provided evidence cues (via the computer) to their initial four questions in the order of their rankings.[7] Evidence direction was manipulated at two levels—confirming and disconfirming evidence. Confirming evidence supported the hypothesis frame the subjects favored, while disconfirming evidence did not. The computer program identified the hypothesis frame the subjects favored and provided them with either consistently confirming evidence or consistently disconfirming evidence.[8] The evidence cues consisted of hypothetical findings of audit tests and procedures performed to address the questions ranked by each subject.[9] The evidence cues were developed by the authors and were reviewed by academics and professional auditors for realism.

After evaluating the evidence cues, the subjects were asked to indicate the hypothesis frame they now favored and to assign a likelihood assessment to that belief. The subjects were then asked to select any of the 12 remaining questions from the menu that they felt would help them identify the source of the ratio fluctuations, although they were free to seek no additional information. Upon verifying their selections, they were asked to rank-order them.

---

[5] This dichotomization into environmental (nonerror) and error frames is consistent with the methodology utilized by Koonce (1992) in an analytical review context. A combination of the two frames was not offered, so as to make the auditors' initial beliefs as strong as possible.

[6] The questions, which are listed in table 5, were randomized, and two versions of the menu were distributed randomly among the subjects. Statistical tests revealed no evidence of order effects; thus, the subjects' responses were pooled for purposes of analysis.

[7] Prior studies used between two and six types of evidence (Asare 1992; Ashton and Ashton 1988, 1990; Einhorn and Hogarth 1985; Tubbs et al. 1990). Half of the subjects evaluated the evidence cues simultaneously and the other half sequentially. This manipulation had no effect on belief revision or evidence search; thus, it was dropped from the discussion.

[8] The use of mixed data (both confirming and disconfirming evidence) has usually been utilized to investigate recency and primacy effects. Although this is an area worthy of audit research, it was considered beyond the scope of the present study.

[9] Evidence direction was randomly distributed across experience levels; 83 subjects received confirming evidence and 83 received disconfirming evidence. Fifty-five of the 83 subjects who evaluated disconfirming evidence changed hypothesis frames. The conclusiveness (or strength) of the audit evidence was controlled for by providing direct answers to the selected audit questions (i.e., the answers to the error questions unambiguously indicated the presence or lack of error, and the feedback to the environmental questions did or did not acknowledge change in the client's policies or competitive environment).

III. Analysis and Results

Belief Revision and Evidence Evaluation

A three-factor (2×2×3) MANCOVA was used to test the hypotheses about belief revision.[10] The primary independent variables in hypotheses H1 and H2, respectively, were hypothesis frame (environmental vs. error) and evidence direction (confirming vs. disconfirming). To reduce the experimental error and to control for any confounding effects, two control variables were included in the model. First, because studies have suggested that experienced auditors may evaluate audit evidence differently from those with less experience (Bonner 1990; Gibbins 1984; Libby and Frederick 1990), audit work experience was included as a blocking variable by dividing the subjects into three experience levels—staff, seniors, and administrators (partners/managers).[11] Second, Church's (1991) findings suggest that auditors with high likelihood assessments may be more likely to downplay disconfirming evidence. Furthermore, Ashton and Ashton (1988) showed that scale effects could possibly influence the magnitude of auditors' belief revisions. Thus, each subject's initial likelihood assessment was included as a covariable.

Belief revision, the dependent variable, was measured two ways. The first measure was based on an "absolute" measurement scale. The absolute scale captured the difference between the likelihood score for a subject's initial hypothesis from (L_0) and the likelihood score for the hypothesis frame favored after the evidence cues had been evaluated (L_1). The absolute scale concentrates on a subject's movement to a final anchor position. Since the subjects were operating in a two-state action space, the dependent variable had to consider the hypothesis frame favored as well as its associated likelihood assessment. For subjects who evaluated either confirming or disconfirming evidence and still favored their initial hypothesis frame, belief revisions were calculated as $L_1 - L_0$. As pointed out earlier, the two hypothesis frames are mutually exclusive and exhaustive; thus, it can be argued that a 70 percent belief in one hypothesis frame implies a 30 percent belief in the other.[12] In this two-state action space, for subjects who evaluated disconfirming evidence and then changed to favor the alternative hypothesis frame, belief revision was calculated as $(100 - L_1) - L_0$.

---

[10] Since both belief revision and evidence search had multiple measures, multivariate analysis of variance was utilized because a series of univariate *F*-tests performed on dependent variables can result in an inflated Type I error rate and ignore any possible correlation among the dependent variables, which could provide the strongest evidence of reliable group differences (Hair et al. 1986, 157). Separate ANCOVA models run on the dependent measures produced nearly identical results. Chi-square tests performed at the $p < 0.05$ level revealed that the assumption of homogenous variances was not violated in any of the statistical models used to analyze the data in this study. Insignificant interaction effects were dropped from the analysis.

[11] These categorizations of experience are supported by the findings of Abdolmohammadi and Wright (1987), which suggest that the effect of experience on audit judgment is not a continuous function. As an alternative, the statistical models were run by treating years of audit experience as a covariate. The effect was insignificant in each case.

[12] Approximately 87 percent of the subjects had initial likelihood assessments greater than 50 percent while only 13 percent of the subjects formed initial likelihood assessments equal to 50 percent. These results support the contention that 50 percent may be too low as an initial likelihood assessment for most auditors. It could be argued that the subjects who assigned a likelihood assessment of 50 percent were indifferent between the two hypothesis frames. As a result, MANCOVAs of the belief revision and search phases excluded those subjects who initially indicated a 50 percent likelihood assessment. The results were essentially unchanged. For example, the primary multivariate effects of evidence direction and the three-way interaction (hypothesis frame × evidence direction × likelihood assessment) remained highly significant ($p < 0.0001$). Hypothesis frame remained highly significant ($p < 0.0001$) for the raw and the ranked search scores. The mean scores for evidence search reported subsequently were only slightly changed.

Table 1
Analysis of Belief Revisions

Panel A. Multivariate Results:

Independent Variable	*F-Value*[a]	*p-Value*
Hypothesis Frame (HF)	5.39	.0055
Evidence Direction (ED)	22.05	.0001
Experience Level	0.52	.7245
Likelihood Assessment (LA)	3.05	.0502
HF × ED × LA	18.35	.0001

Panel B. Univariate Results:

Independent Variable	*df*	SS	*MS*	*F-Value*	*p-Value*
Absolute Scale					
Hypothesis Frame	1	2154.93	2154.93	9.74	.0021
Evidence Direction	1	1879.13	1879.13	8.49	.0041
Experience Level	2	125.81	62.91	0.28	.7529
Likelihood Assessment	1	1054.17	1054.17	4.77	.0305
HF × ED × LA	3	7674.89	2558.30	11.56	.0001
Proportional Scale					
Hypothesis Frame	1	0.72	0.72	10.62	.0014
Evidence Direction	1	0.00	0.00	0.01	.9240
Experience Level	2	0.10	0.05	0.68	.5063
Likelihood Assessment	1	0.41	0.41	6.17	.0143
HF × ED × LA	3	0.83	0.28	4.13	.0075

[a] Wilk's criterion.

The second measure was based on a "proportional" scale designed to moderate scale effects. It captured the relative change in the subjects' reactions to the evidence cues by considering the maximum revision that was available to each subject, given the range constraints of the experiment. The available revision range for each subject was based on the initial likelihood assessment and the direction of evidence evaluated. For example, if a subject's initial likelihood score was 70 percent and the subject evaluated confirming evidence, a 30-point range was available for upward revision. Conversely, if the subject evaluated disconfirming evidence, a 70-point range was available for downward revision. Therefore, for subjects who evaluated confirming or disconfirming evidence and who did not switch hypothesis frames, belief revisions were calculated as $(L_1 - L_0)/(100 - L_0)$. For subjects who changed their hypothesis frame after evaluating disconfirming evidence, belief revision was calculated as $[(100 - L_1) - L_0]/L_0$.[13]

Hypothesis H1 predicted that the magnitude of the belief revision would not be affected by hypothesis frame. The MANCOVA results for the dependent measures are presented in panel A of table 1. Hypothesis frame was significant ($p < 0.0055$). Univari-

---

[13] In conformance with Tubbs et al. (1990), Ashton and Ashton (1988, 1990), and Bamber et al. (1991), only the reactions of the subjects who understood the directional implications of the audit evidence were considered. Of the 175 total subjects who participated in the study, only nine (5 percent) made likelihood revisions in the wrong direction (i.e., four subjects revised downward after receiving confirming evidence and five subjects revised upward after receiving disconfirming evidence). These subjects were dropped from the analysis leaving 166 useable observations.

Table 2
Mean Belief Revisions

	Hypothesis Frame		
Evidence Direction	*Environmental*	*Error*	*Total*
Absolute Scale			
Confirming Evidence	14.31	17.85	15.17
Disconfirming Evidence	27.54	32.58	29.00
Total	20.73	25.70	22.08
Proportional Scale			
Confirming Evidence	0.487	0.552	0.503
Disconfirming Evidence	0.394	0.475	0.417
Total	0.442	0.511	0.476

ate analyses to determine which dependent variables were affected by hypothesis frame (Bray and Maxwell 1986, 40) are shown in panel B. Hypothesis frame was significant for both the absolute measure ($p<0.0021$) and the proportional measure ($p<0.0014$). Thus, Hypothesis H1 was rejected.

The subjects' mean likelihood revisions by hypothesis frame and evidence direction are presented in table 2. Since the magnitude of the revisions was the major issue of interest, the mean disconfirming revisions were treated as positive values in the statistical procedures. The mean error revisions (absolute, 25.70; proportional, 0.511) were greater than the mean environmental revisions (absolute, 20.73; proportional, 0.442).

Hypothesis H2 predicted that the direction of the evidence evaluated would have no effect on the magnitude of the subjects' belief revisions. As shown in panel A of table 1, evidence direction was highly significant ($p<0.0001$). However, the univariate analyses (panel B) reveal that the two measures of belief revision did not provide consistent results. For the absolute scale, evidence direction was significant ($p<0.0041$), which supports rejection of hypothesis H2. However, for the proportional scale, hypothesis H2 was not rejected as the evidence direction effect was highly insignificant ($p<0.9240$). As shown in table 2, the absolute measure of the revisions made by subjects who evaluated disconfirming evidence (29.00) was almost twice as large as the revisions of the subjects who evaluated confirming evidence (15.17). In contrast, proportional measures of belief revisions resulting from the evaluation of confirming evidence (0.503) and disconfirming evidence (0.417) were in the opposite direction. However, the two proportional means were not significantly different.

Although experience level was not found to have a significant effect on belief revisions, likelihood assessment was marginally significant in the multivariate analysis at conventional levels of significance ($p<0.0502$, table 1). The univariate analyses revealed that the likelihood covariate was significant for both the absolute measure ($p<0.0305$) and the proportional measure ($p<0.0143$). These results imply that the extremity of the subjects' belief (i.e., the likelihood that the hypothesis frame represented the true state) affected the magnitude of the belief revision. A significant three-way interaction among hypothesis frame, evidence direction, and likelihood assessment was evident in both the multivariate ($p<0.0001$) and univariate analyses (absolute

Table 3
Belief Revision Slopes

	Hypothesis Frame	
Evidence Direction	*Environmental*	*Error*
Absolute Scale		
Confirming Evidence	−0.538*	−0.035
Disconfirming Evidence	0.156	1.406*
Proportional Scale		
Confirming Evidence	−0.002	0.014**
Disconfirming Evidence	−0.004	0.015*

* Significant at $p<0.01$.
** Significant at $p<0.10$.

$p<0.0001$, proportional $p<0.0075$). This implies that subjects reacted differently to confirming and disconfirming evidence, depending on the hypothesis frame favored and the extremity of their likelihood assessment. Hays (1981) notes that significant interactions provide interesting qualifications of the main effects. Furthermore, according to Neter et al. (1985), if the covariate interacts with any of the other independent variables, separate regression lines should be estimated for each group. Thus, the interaction was further examined by calculating separate regression coefficients for each of the four combinations of hypothesis frame and evidence direction. Table 3 shows the parameter estimates (i.e., regression slopes) for the belief revisions with absolute and proportional dependent measures. The eight regression lines are graphically illustrated in figure 1.

The negative coefficients for confirming evidence for three of the four regressions indicate that an inverse relationship existed between the subjects' upward revisions and their initial likelihood assessments. In other words, the greater the initial likelihood assessment, the smaller the auditors' belief revisions. This behavior is consistent with hypothesis H4. This inverse relationship was significant for auditors who initially favored the environmental hypothesis frame under the absolute scale (slope of −0.538). The positive slope parameters for three of the four regressions for disconfirming evidence indicate that downward revisions increased with the extremity of the subjects' initial likelihood assessment. The indirect relationship in this case was significant for auditors who favored the error hypothesis (absolute slope of 1.406, proportional slope of 0.015).

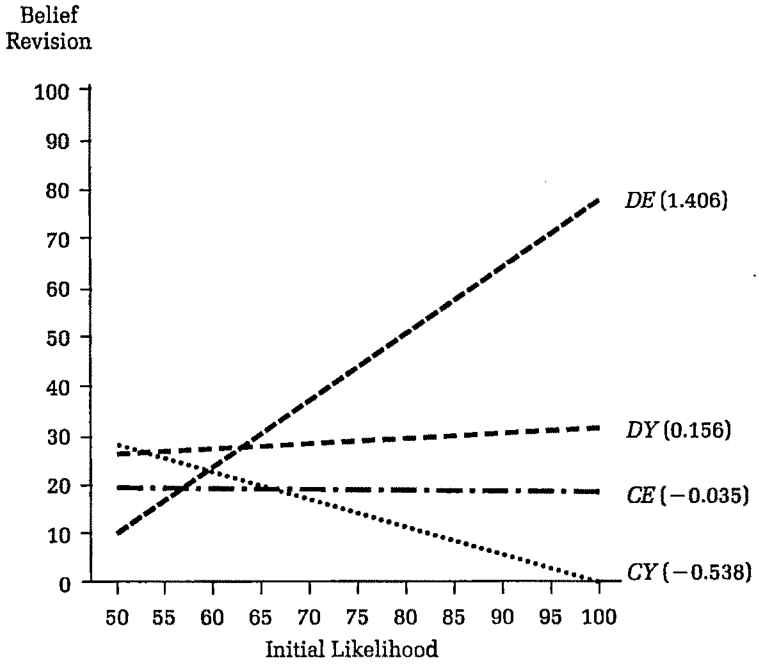
The preceding relationships did not hold under the proportional scale for auditors who evaluated confirming evidence and favored the error frame or for auditors who evaluated disconfirming evidence and favored the environmental frame. The slope for error frame (0.014) was marginally significant ($p<0.0945$), which suggests that, for error-framed subjects, the greater their initial likelihood assessment, the greater their relative belief revision after evaluating confirming evidence.

Evidence Search

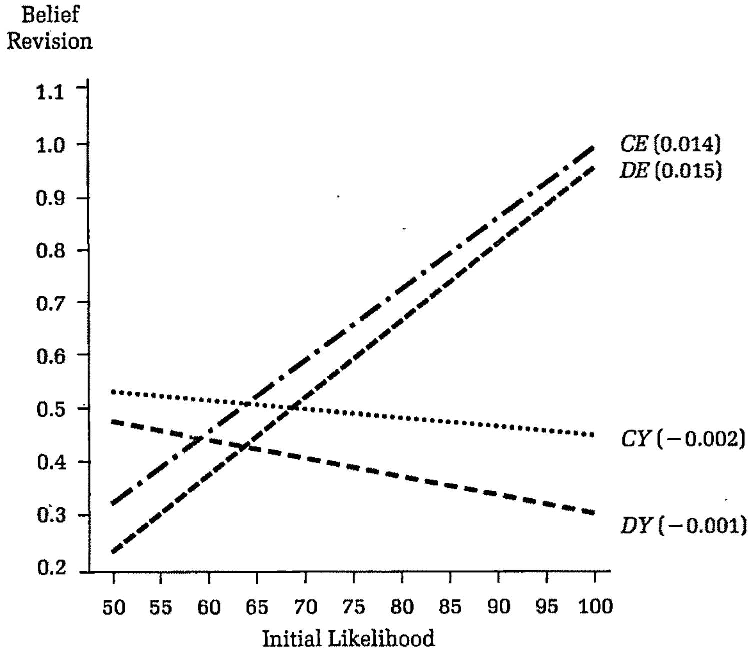
A three-factor (2 × 2 × 3) MANCOVA model was used to test the hypotheses related to evidence search. In conformance with the analysis for belief revision, hypothesis

Figure 1
Belief Revision Slopes By Evidence Direction and Hypothesis Frame

Absolute Scale



Proportional Scale



Note: The Treatment Combinations are:

— · — *CE*= confirming evidence and error frame;

frame (revised) and evidence direction were used to test Hypotheses H3 and H4. Similarly, audit experience level was included as a blocking variable and likelihood assessment (revised) was included as a covariable to reduce the experimental error and to control for confounding effects.

Two dependent measures were used to assess the subjects' search behavior. The first, labeled as a "raw" search score, measured whether one type of evidence (environmental vs. error) was sought more often than the other. Recall that subjects were provided evidence according to their top four ranked questions selected from a list of eight environmental and eight error questions. Consequently, the remaining questions would not be equally distributed between the two types of evidence for subjects who did not initially rank two of each type of evidence. In turn, the probability of choosing either an environmental or an error question at random would no longer be 50 percent but, rather, equal to the remaining number of environmental or error questions divided by 12. To compensate for this potential bias, each question selected was weighted by the reciprocal of the probability that one question of its evidence type would be selected. For example, if the 12 remaining questions consisted of eight environmental questions and four error questions, each environmental question was assigned a weight of $12/8$ and each error question was assigned a weight of $12/4$. An equal number of questions remaining for each type resulted in a weighting factor of $12/6$ for each type of question. (See the appendix for a chart of all ten possible outcomes for the initial search.) The raw search scores were created by multiplying a question's weighting factor by 1 when the question was consistent with a subject's revised hypothesis frame and by -1 when it was inconsistent with the revised hypothesis frame. These values were then summed to form each subject's raw score.

The second dependent variable was a "ranked" search score designed to place more emphasis on the evidence cues that the subjects chose to examine first. The rank assigned to a question was based on the number of questions selected by a subject (0 through 12), where the highest rank was assigned to the top-ranked question, and correspondingly lower ranks were assigned to the other questions. For example, if the auditor selected three questions, the ranks assigned to each question were 3, 2, and 1 for the first, second, and third rankings, respectively. Questions that agreed with a subject's hypothesis frame were assigned positive ranks, and negative ranks otherwise. As in the raw search score, the ranks were adjusted by the weighting factor described above, summed, and standardized by dividing by the number of questions selected. The latter adjustment compensated for the fact that subjects were not required to select an equal number of evidence cues.[14]

Hypothesis H3 predicted that the type of evidence selected would not be influenced by the auditor's hypothesis frame. Panel A of table 4 shows the overall MANCOVA results for the evidence search measures, and panel B shows the univariate results. Both the multivariate and univariate analyses produced significant hypothesis frame effects ($p<0.0001$ in all cases). Thus, hypothesis H3 was rejected, as the auditors' hypothesis frame significantly influenced the type of evidence selected.

---

[14] For example, the raw score for a subject who favored the environmental frame and selected four environmental questions and two error questions when the menu consisted of seven environmental questions and five error questions would be $2.06=(12/7+12/7+12/7+12/7-12/5-12/5)$. If the first four highest ranked questions were the environmental questions, the ranked score would be $23.66=\{[(6*12/7)+(5*12/7)+(4*12/7)+(3*12/7)-(2*12/5)-(1*12/5)]\}$.

Table 4
Analysis of Evidence Search

Panel A. Multivariate Results:

Independent Variable	*F-Value*[a]	*p-Value*
Hypothesis Frame	19.67	.0001
Evidence Direction	0.75	.4752
Experience Level	0.43	.7836
Likelihood Assessment	3.15	.0457

Panel B. Univariate Results:

Independent Variable	*df*	*SS*	*MS*	*F-Value*	*p-Value*
Raw Search Score					
Hypothesis Frame	1	518.60	518.60	26.93	.0001
Evidence Direction	1	1.20	1.20	0.06	.8034
Experience Level	2	3.48	1.74	0.09	.9136
Likelihood Assessment	1	2.85	2.85	0.15	.7009
Ranked Search Score					
Hypothesis Frame	1	569.86	569.86	38.18	.0001
Evidence Direction	1	15.79	15.79	1.06	.3052
Experience Level	2	5.61	2.81	0.19	.8289
Likelihood Assessment	1	33.39	33.39	2.24	.1367

[a] Wilk's criterion.

The specific effect of hypothesis frame on the auditors' evidence search is shown by the mean evidence search scores presented below by hypothesis frame.

Hypothesis Frame	*Raw Score*	*Ranked Score*
Environmental	−0.86	−0.96
Error	3.06	2.97

For both the raw and ranked search measures the greater the positive scores, the greater the confirmatory behavior. Negative search scores suggest disconfirmatory behavior and a zero search score indicates a perfectly balanced search strategy. Confidence intervals constructed for each of the means revealed that three of the means were different from zero at the 99 percent significant level, and one (−0.86) was different from zero at the 95 percent significance level. The mean scores indicate that the subjects who favored the material error frame after evaluating the first four evidence cues designed their continued evidence search to favor the uncovering of material error. However, the subjects who favored the environmental frame did not place more weight on environmental evidence.

Hypothesis H4 predicted that the direction of evidence initially evaluated would have no effect on the type of additional evidence selected. Panel A of table 4 shows that evidence direction was not significant ($p<0.4752$). Thus, hypothesis H4 was not rejected, which indicates that the search scores were not influenced by whether subjects evaluated confirming or disconfirming evidence.

The multivariate and univariate analyses revealed that neither experience level nor likelihood assessment had any confounding effects on the subjects' search behavior. In addition, ancillary analysis indicated that the mean number of additional evidence cues requested (approximately 5.8) was not affected by hypothesis frame, evidence direction, or experience level.

IV. Summary Discussion and Implications

Belief Revision

The results suggest that different experimental methods (i.e., allowing subjects to generate their own initial likelihood assessments vs. providing them with a given anchor) and different dependent measurements of belief revision (absolute vs. proportional) may provide possible explanations for the mixed results in research on auditors' reactions to confirming and disconfirming evidence. In the first phase of the experiment, the evidence suggests that an experimental condition that artificially presets subjects' initial anchor at 50 percent is not actually reflective of most auditors' initial beliefs. If subjects adjust the experimental anchor after considering their prior experiences, the measurement of belief revisions following an experimental manipulation may be biased.

In this study, the subjects' belief revisions were measured in two ways. When the absolute scale was used, the results were consistent with those of Ashton and Ashton (1988, 1990). That is, the auditors were more responsive to disconfirming than confirming evidence. However, when the proportional scale was used, the analysis revealed that the subjects were equally sensitive to confirming and disconfirming evidence (table 2).

The results of this study also reveal that the frame of the hypothesis tested can have a significant effect on how auditors react to confirming and disconfirming evidence. Auditors who began the audit task by favoring the error frame reacted more to both confirming and disconfirming evidence than did auditors who began by favoring the environmental frame. This suggests that auditors with an error frame are likely to be more sensitive to the receipt of audit evidence, whether it is confirmatory or disconfirmatory, than auditors who are operating with an environmental frame.

The belief revisions exhibited by the subjects in this study suggest potential effects on audit efficiency and effectiveness. Subjects who initially favored the error frame reacted strongly to disconfirming evidence. After evaluating evidence that did not support the notion that a material error was likely, these subjects moved rather quickly toward favoring the environmental frame as the more likely explanation for the ratio fluctuations. As noted in figure 1, the absolute change for subjects evaluating disconfirming evidence under the error frame increased sharply as the initial likelihood assessment increased, which made the relative change under the proportional scale an increasing function. These results are consistent with the results of Ashton and Ashton (1988). In their study, the most dramatic change on an absolute and relative basis occurred when the inital error rate was set at 80 percent and the subjects evaluated disconfirming evidence (absolute change was 36.1 percent and the relative change was 45.1 percent).

The behavior of the subjects who evaluated disconfirming evidence and favored the error frame is not consistent with Church's (1991) contention that the more committed auditors are to a hypothesis, the more likely they are to exhibit confirmatory behavior. To conform with confirmatory bias, these auditors should have downplayed the disconfirming evidence. However, the reverse was true. The behavior also is inconsistent with the notion that auditors should be highly concerned about the possible existence of material errors.

One possible explanation of why the error-framed subjects reacted so acutely to disconfirming evidence is that, even though auditors may initially feel that material errors are likely they also are aware that the occurrence of material errors is not that high (Libby and Frederick 1990). Thus, after finding no evidence of material errors in the most likely areas, the subjects switched their focus to environmental conditions. However, for error-framed subjects who evaluated confirming evidence, the proportional scale slope in figure 1 suggests that the stronger the subjects' belief that a material error occurred, the greater the relative upward adjustment.

With respect to the subjects who began the audit judgment process favoring the environmental frame, the results indicate a tendency to downplay disconfirmatory evidence. This behavior was prevalent among these subjects regardless of the extremity of their belief. As illustrated by the flat slope under the absolute scale in figure 1, the downward belief revision after the evaluation of audit evidence (about 27 percentage points) was affected by the auditor's initial likelihood assessment. When the change was measured as a percentage of the potential change, a negative slope occurred. Behavior of this nature is not consistent with the notion of conservative bias or error-sensitive behavior. It suggests that auditors may have a tendency to underweight potentially important error-related evidence when they are not focused on detecting material errors. This may happen because confirmation bias mitigates their sensitivity to error evidence.

Although it is important that auditors demonstrate sensitivity to material-error information, the over- or underreaction to any audit evidence could lead to premature closure with an incorrect hypothesis or, at a minimum, audit inefficiences. Furthermore, auditors who consistently encounter confirming evidence early in an audit investigation must be careful not to close on a hypothesis prematurely without considering alternative hypotheses. The indication in this study was that the auditors did not demonstrate a propensity for premature closure. Specifically, the subjects who initially evaluated confirming evidence did not ask for less additional information than the subjects who evaluated disconfirming evidence.[15]

Experience level had no significant effect on how the auditors evaluated the audit evidence. Staff auditors made judgments similar to those of more experienced auditors concerning the effect and importance of an audit cue. Since staff auditors conduct a substantial amount of compliance and substantive testing, these results are favorable to the profession. The lack of an experience effect is consistent with the results of Bamber et al. (1991), but are inconsistent with the results of Bedard and Biggs (1991), Bonner (1990), Libby and Frederick (1990), and Kaplan and Reckers (1989). The inconsistent results may be explained by the different tasks examined in these studies. In Bedard and Biggs (1991), Bonner (1990), and Libby and Frederick (1990), the auditors essentially were graded on the number of correct explanations of a possible outcome. In the current study and in the study by Bamber et al. (1991), the focus was on whether the auditors had different belief revisions after evaluating evidence, rather than on generating plausible explanations.

---

[15] A potential explanation for this behavior is that subjects viewed the additional audit information as costless (cost/benefit considerations are difficult to operationalize in an experimental task). With this logic, the subjects should have requested all the information available (i.e., selected all 12 questions). However, only 9 percent of the subjects selected ten or more questions, which suggests some awareness of cost/benefit considerations by the subjects.

Evidence Search

The hypothesis frame (revised) favored by the subjects had a significant effect on their continued search for evidence. The mean scores for evidence search appear to indicate that the subjects who favored the error frame acted in a confirmatory manner (i.e., selected more error questions and ranked the receipt of error evidence as more important than environmental evidence). In contrast, the subjects who favored the environmental frame acted in a disconfirmatory manner (i.e., selected more error questions and gave higher rankings to error questions).

The professional and legal environment of auditing may explain why auditors behave differently from subjects in the psychological literature. The auditor's primary risk is that a material error in a client's financial statements will go undetected. The auditors in this study appeared to conduct their investigation with the objective of reducing this risk by designing their evidence search to emphasize the uncovering of potential material errors, regardless of the hypothesis frame they favored. This finding is consistent with Smith and Kida's (1991) contention that auditors' behavior in most cases will be conservatively biased toward uncovering error-related evidence.

Overall, the results of this study suggest that confirmatory bias may have enhanced the conservative bias of auditors favoring the error frame and may have weakened sensitivity to error evidence when the environmental frame was favored. The continued search of many subjects was fairly balanced, even after they had received all confirming evidence. The employment of balanced search strategies could be considered favorable behavior in most circumstances, as it indicates a resistance to premature closure. SAS No. 31 (AICPA 1980) stresses that auditors need to be thorough in their search for evidential matter and need to consider all evidence regardless of whether it appears to corroborate or contradict the assertions in the financial statements.

Nevertheless, strict adherence to conservative or balanced search strategies can result in audit inefficiencies. Behavior of this nature may have been suggested, as the extremity of the subjects' beliefs had no effect on their scores for continued search. This behavior could reflect the difficulty auditors have in balancing the profession's competing objectives of efficiency and effectiveness.

A least-effort strategy or a diagnostically driven strategy are two alternatives that could have affected the auditors' evidence search. Given equal diagnosticity, a least-effort strategy implies that subjects initially select audit procedures that are the easiest or least time-consuming to perform. Typically, environmental evidence is less time-consuming to analyze than error-related evidence. A diagnostically driven strategy implies that auditors prefer those audit procedures that are the most diagnostic when the cost or difficulty of the procedures prescribed by each evidence cue is equal. It could be argued that, in general, error evidence is often considered more diagnostic than environmental evidence because it is more likely to be obtained from an objective examination of supporting documents. Because environmental evidence often takes the form of verbal and written representations by the client, it may be perceived as less reliable.

If the subjects' search behavior was driven by a least-effort or a diagnostic strategy, it should not be greatly affected by hypothesis frame. However, table 5 clearly shows that hypothesis frame influenced the subjects' search. Note that the direction of all the significant differences between the percentage of subjects selecting an audit question is consistent with the hypothesis frame favored. Furthermore, of the four most frequently selected questions in the initial search, two were environmental and two were error. The same holds for the four least selected questions. Of the eight most overall requested questions, five were error and three were environmental. Therefore, when hypothesis frame is considered, it does not appear that a least-effort or diagnostic-driven strategy had any major influence on the subjects' evidence behavior.

Positive implications may be drawn from the fact that neither evidence direction nor experience level had a significant effect on the auditors' continued search. It appears that the feedback provided from their evidence selection helped mitigate the confirmatory tendencies of the staff auditors noted in the first phase of the experiment (see McMillan and White 1992).

The usual limitations associated with field experiments are applicable to this study. Generalizability is always a concern; however, a wide range of subjects from several different firms participated in this study, and analyses of the subjects' demographic data (i.e., firm, education, gender) had no effect on the overall results.[16] Though the experiment involved no explicit incentives, the subjects' cooperation had been requested by their senior audit partner, and comments from the participants indicated that they found the task interesting.

The results of this study also may be limited by the methodology employed and by the way the variables were measured. First, the two-state action space used in this study and the simultaneous selection of audit evidence may be overly simplistic. Although it is not uncommon to receive several evidence cues from the client at the same time during the early phase of an audit investigation, different results may have occurred in this study had the subjects been able to choose each evidence cue sequentially after first evaluating the feedback of the previously selected evidence cue. Second, measurements of belief revision other than with the absolute and proportionate scales may have resulted in different outcomes. Third, the search process may have been influenced by the diagnostic content of the 16 evidence cues. Although the competence level of each cue was not measured, the frequency of cue selection was evenly distributed over a large number of cues. This suggests that many cues had similar diagnostic content. Future research should address these issues to further reconcile the different results among the studies in this area. Lastly, the lack of a unified model that addresses auditor decision making during the multiple phases of an audit contributes to inconsistent results.

---

[16] The mean time taken to complete the entire audit task was approximately 26 minutes. Cumulative time intervals recorded at several logical points produced no significant results or changes in the MANCOVA for belief revision. However, the MANCOVA for evidence search revealed that the time spent in evaluating audit feedback significantly affected the subjects' continued search (Wilks $F=4.70$, $p<0.0036$; raw univariate $p<0.008$ and ranked univariate $p<0.008$). The coefficient for the evaluation time period covariate was negative, which indicates that the longer the subjects spent evaluating the evidence, the less was the tendency to exhibit confirmatory behavior. This could suggest that greater "effort" reduced confirmatory tendencies. An ANCOVA analysis performed with evaluation time as the dependent variable revealed that none of the independent variables had a significant effect on the time subjects spent evaluating the audit evidence they requested (model $F=1.74$, $p<0.1141$).

Table 5
Percentage of Subjects Selecting Audit Question by Hypothesis Frame

	Initial Search			Continued Search			Combined Search	
Audit Question	Environmental Frame	Error Frame	Total	Environmental Frame	Error Frame	Total	Total	Rank
A. Has management instituted more liberal terms for credit customers?	54*	20*	45	36	30	34	78	2.5
B. Have some of next periods' sales been improperly recorded this period?	35*	51*	39	39	41	40	79	1
C. Were bad debt expenses not properly recorded this period?	19	24	20	47	47	47	67	7
D. Has the rate of dividend payment been reduced this year?	7	7	7	29	25	28	34	16
E. Were accrued expenses not properly recorded this period?	34*	53*	39	34	47	39	78	2.5
F. Have some expense items been improperly recorded as inventory?	22*	42*	28	40	49	44	72	5
G. Were purchase returns not properly recorded this period?	11	8	10	35*	56*	43	53	13
H. Has the firm recently liquidated any significant asset holdings?	33	22	30	33	29	31	61	8
I. Has the firm expanded operations of its retailing division?	23	16	21	33*	17*	27	48	14.5
J. Has management instituted new inventory policies which led to a reduction of raw materials on hand and lower current liabilities?	24	16	22	36	30	34	55	11.5
K. Have marketable securities not been properly adjusted to lower of cost or market?	18	22	19	39	31	36	55	11.5
L. Has the firm accelerated its payments to suppliers?	15	9	13	39	28	35	48	14.5
M. Has inventory turnover slowed?	26	29	27	34	27	31	58	9
N. Were payments on account made next period improperly recorded this period?	12	11	12	39*	55*	45	57	10

Continued

Table 5—*Continued*

Audit Question	*Initial Search*			*Continued Search*			*Combined Search*	
	Environmental Frame	*Error Frame*	*Total*	*Environmental Frame*	*Error Frame*	*Total*	*Total*	*Rank*
O. Have purchases made and received this period not been recorded or improperly recorded next period?	22	33	25	43	48	45	70	6
P. Have proceeds from major new loans been obtained this period?	45	36	43	27	36	31	73	4

Note: Environmental questions are A, D, H, I, J, L, M, and P; the error questions are B, C, E, F, G, K, N, and O. The most frequently selected question was ranked 1, the next 2, and so forth. In the initial search, there were 121 subjects in the environmental frame and 45 in the error frame. In the continued search, there were 102 subjects in the environmental frame and 64 in the error frame.

* Denotes a statistical difference at $p<.05$.

Implications for Future Research

The results of this study suggest that confirmatory behavior may not prevail or dominate in the audit judgment process. In fact, the subjects exhibited behavior consistent with an error-bias attitude. Many may consider error-bias behavior favorable because it indicates a responsiveness to professional skepticism and a demonstration of conservatism. However, the actual strength and role of this decision bias needs further investigation because of its potential effects on the efficiency and effectiveness of audit decision making.

Further research is especially needed to investigate the possibility suggested by this study that auditors who begin the audit judgment process feeling that material errors are unlikely may be inclined to downplay evidence that indicates that material errors may exist. Just as potentially dangerous is the implication that auditors who are initially inclined to believe that material errors may exist have a tendency to reverse this position after an initial review of evidence that does not confirm that belief. It would be interesting to see whether such factors as time constraints or performance evaluations might play a role in this context.

The balanced search strategies employed by the subjects in this study signify that auditors are inclined to look for evidence that has diagnostic content (i.e., evidence that allows them to distinguish between alternatives) and relevant information. Evaluating the diagnostic content allotted to confirming and disconfirming evidence and examining how auditors handle irrelevant evidence could be fruitful areas of future research. Furthermore, it is possible that experience has a significant role to play in the audit judgment process when auditors encounter irrelevant as well as relevant information.

Appendix

The following chart presents all ten possible outcomes for the initial search on the basis of the hypothesis frame favored and the four ranked questions selected. In all cases, the maximum score is +12 and the minimum score is −12; Y denotes environmental questions and E denotes error questions.

Hypothesis Frame	*Initial Rankings*	*Questions Remaining*	*Weighting Factor*	
			Environmental	*Error*
Environmental	Y=4 E=0	Y=4 E=8	Y*(12/4)	E*(12/8)
Error	Y=0 E=4	Y=8 E=4	Y*(12/8)	E*(12/4)
Environmental	Y=3 E=1	Y=5 E=7	Y*(12/5)	E*(12/7)
Error	Y=1 E=3	Y=7 E=5	Y*(12/7)	E*(12/5)
Environmental	Y=2 E=2	Y=6 E=6	Y*(12/6)	E*(12/6)
Error	Y=2 E=2	Y=6 E=6	Y*(12/6)	E*(12/6)
Environmental	Y=1 E=3	Y=7 E=5	Y*(12/7)	E*(12/5)
Error	Y=3 E=1	Y=5 E=7	Y*(12/5)	E*(12/7)
Environmental	Y=0 E=4	Y=8 E=4	Y*(12/8)	E*(12/4)
Error	Y=4 E=0	Y=4 E=8	Y*(12/4)	E*(12/8)

References

Abdolmohammadi, M., and A. Wright. 1987. An examination of the effects of experience and task complexity of judgments. *The Accounting Review* 62 (January): 1–13.

American Institute of Certified Public Accountants. 1980. *Evidential Matter: Statement on Auditing Standards No. 31*. New York: AICPA.

——. 1988. *Analytical Procedures: Statement on Auditing Standards No. 56*. New York: AICPA.

Asare, S.K. 1992. The auditor's going-concern opinion decision: Interaction of task variables and the sequential processing of evidence. *The Accounting Review* 67 (April): 379–93.

Ashton, R. H., and A. H. Ashton. 1988. Sequential belief revision in auditing. *The Accounting Review* 63 (October): 623–41.

——, and ——. 1990. Evidence-responsiveness in professional judgment: Effects of positive versus negative evidence and presentation mode. *Organizational Behavior and Human Decision Processes* 46 (June): 1–19.

Bamber, E. M., M. R. Tubbs, G. Gaeth, and R. J. Ramsay. 1991. Auditors' attitudes to confirming and disconfirming evidence in belief revision. Working paper, University of Georgia, Athens.

Bedard, J. C., and S. F. Biggs. 1991. Pattern recognition, hypothesis generation, and auditor performance in analytical review. *The Accounting Review* 66 (July): 622–42.

Bonner, S. E. 1990. Experience effects in auditing: The role of task-specific knowledge. *The Accounting Review* 65 (January): 72–92.

Bray, J. H., and S. E. Maxwell. 1986. *Multivariate Analysis of Variance*. Sage University Paper Series on Quantitative Applications in the Social Sciences, #07–054. Beverly Hills, CA: Sage.

Butler, S. 1986. Anchoring in the judgmental evaluation of audit samples. *The Accounting Review* 61 (January): 101–11.

Church, B. K. 1990. Auditors' use of confirmatory processes. *Journal of Accounting Literature* 9: 81–112.

——. 1991. An examination of the effect that commitment to a hypothesis has on auditors' evaluations of confirming and disconfirming evidence. *Contemporary Accounting Research* 7 (Spring): 513–24.

Einhorn, H. J., and R. M. Hogarth. 1985. A contrast/surprise model for updating beliefs. Unpublished manuscript, Graduate School of Business, University of Chicago.

Gibbins, M. 1984. Propositions about the psychology of professional judgment in public accounting. *Journal of Accounting Research* 22 (Spring): 103–25.

Gorman, M. 1986. How the possibility of error effects falsification on a task that models scientific problem solving. *British Journal of Psychology* 77 (February): 85–96.

Hair, J. F., R. E. Anderson, and R. L. Tatham, eds. 1986. *Multivariate Data Analysis*. New York: Macmillan.

Hays, W. J., ed. 1981. *Statistics*. New York: Holt, Rinehart & Winston.

Kaplan, S. E., and P. M. J. Reckers. 1989. An examination of information search during initial audit planning. *Accounting, Organizations and Society* 14 (May/June): 539–50.

Kida, T. 1984. The impact of hypothesis-testing strategies on auditors' use of judgment data. *Journal of Accounting Research* 22 (Spring): 332–40.

Klayman, J., and Y. Ha. 1987. Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review* 96 (April): 211–28.

Koonce, L. 1992. Explanation and counterexplanation during audit analytical review. *The Accounting Review* 67 (January): 59–76.

Libby, R. 1985. Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research* 23 (Autumn): 648–67.

——, and D. M. Frederick. 1990. Experience and the ability to explain audit findings. *Journal ofAccounting Research* 28 (Autumn): 348–67.

Lord, C., L. Ross, and M. Lepper. 1979. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology* 37 (November): 2098–2109.

McMillan, J. J., and R. A. White. 1992. A re-examination of auditors' initial evidence search behavior. Working paper, Clemson University, SC.

Neter, J., W. Wasserman, and M. H. Kutner, eds. 1985. *Applied Linear Statistical Models*. Homewood, IL: Richard D. Irwin.

Smith, J. F., and T. Kida. 1991. Heuristics and biases: Expertise and task realism in auditing. *Psychological Bulletin* 111 (May): 472–89.

Snyder, M. 1981. Seek and you shall find: Testing hypotheses about other people. In *Social Cognition: The Ontario Symposium*, Vol. 1, edited by E. T. Higgins, C. P. Herman, and M. P. Aznna. Hillsdale, NJ: Lawrence Erlbaum.

———, and W. B. Swann. 1978. Hypothesis testing processes in social interactions. *Journal of Personality and Social Psychology* 36 (November): 1202–12.

Tubbs, R. M., W. F. Messier, and W. R. Knechel. 1990. Recency effects in the auditor's belief-revision process. *The Accounting Review* 65 (April): 452–60.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 466–481

The Effects of Intergroup Competition and Intragroup Cooperation on Slack and Output in a Manufacturing Setting

S. Mark Young
University of Southern California
Joseph Fisher
Indiana University
Tim M. Lindquist
University of Northern Iowa

SYNOPSIS AND INTRODUCTION: Firms are redesigning operations to reduce slack and waste and improve performance (Hoerr 1989; Safizadeh 1991; Walton 1987) and this often involves reorganizing production workers into workgroups to foster cooperation and group participation in setting standards (Hayes et al. 1988; Schonberger 1986). In addition to employing incentive schemes linked to meeting standards, many firms are using bonuses tied to relative performance among groups to develop a spirit of intergroup competition.

Over the past two years, we made several visits to three *Fortune 500* manufacturing firms involved in such changes.[1] The site visits suggested several hypotheses that merited further investigation. Thus, we designed a laboratory experiment to study more systematically what we had observed

---

[1] Since the firms that we visited were all experimenting with the workgroup form of reorganziation, they asked that we reveal neither their names nor details about their businesses in any research projects. The firms produce consumer goods such as electronic instruments, leather goods, and small data storage devices. The hypotheses were formulated and the experiment designed according to a combination of insights gained from the firms, rather than on what any particular firm had implemented.

We would like to thank the KPMG Peat Marwick Foundation; the Amos Tuck School of Business Administration, Dartmouth College; and the University of Colorado, College of Business for financial support and Wayne Wasechek for his research assistance. In particular, we thank Alison Ashton, Bob Ashton, Sarah Bonner, Jon Davis, Jennifer Francis, Jim Frederickson, Reid Hastie, Laureen Maines, Jamie Pratt, Frank Selto, Bill Waller, and S. S. Y. Young for detailed comments on previous drafts of this article and participants in the accounting workshops at Duke University, Indiana University, and the University of Wisconsin for very useful discussions.

Submitted April 1992.
Accepted December 1992.

in the field. This study extends previous research on determinants of slack and performance (e.g., Chow 1983; Chow et al. 1988; Waller and Chow 1985; Young 1985) by: (1) incorporating our observations and the literature on intragroup cooperation and competitive feedback to develop hypotheses, (2) studying workgroups rather than individuals, and (3) using a multiperiod rather than single-period setting.[2]

Results of the experiment show that the type of competitive feedback received by groups affected both their output and slack. Interestingly, when individuals were allowed to cooperate rather than work in isolation, performance actually declined. This latter result was unexpected and was likely the consequence of the particular experimental task.

Key Words: *Manufacturing performance, Slack, Cooperation, Competitive feedback.*

Data Availability: *Inquiries about the data collection materials and data may be addressed to the first author.*

IN section I of the article, we review the relevant literature and develop hypotheses about the effects of intragroup cooperation and competitive feedback. We describe the experimental procedures in section II, report the results of the experiment in section III, and discuss our findings in the final section.

I. Background and Literature Review

Cooperation, Competition, and Task Interdependence

There is an extensive literature on the effects of individual inputs on group outcomes and compensation (Deutsch 1949; Goldman et al. 1977; Okun and DiVesta 1975; Schmitt 1981). Group goals and performance have been studied in intra- and intergroup cooperative and competitive settings (Johnson et al. 1981) with various types of incentive systems (Miller and Hamblin 1963; Steiner 1972). The degree of interdependence among tasks performed by individuals also effects both cooperation and competition (Schmitt 1981). With low interdependence, it is feasible and economical for a single individual to perform the entire task, but highly interdependent tasks involve the efforts of more than one person in creating a final product or outcome.

In the firms that we visited, workgroups were involved with intragroup cooperation, but were also faced with intergroup competition in highly interdependent tasks (assembly line production). All three firms had chosen this particular configuration as they believed it to be the most motivating. The idea was to develop a work culture in which individuals no longer competed against others on the line, but rather cooperated as a group through joint problem solving and team building. To promote even higher performance, each firm (with slight variations) had developed a competitive system in which workgroups were pitted against each other for rewards. The configuration chosen by these firms is quite consistent with the findings of previous experimental research that the greatest level of performance for highly task-interdependent groups occurs when a group is internally cooperative but faces intergroup competition (Okun and Divesta 1975; Schmitt 1981; Tjosvold 1984).

---

[2] Young et al. (1988) studied performance in groups but did not manipulate group level variables; Chow et al. (1991) conducted a study over several periods but with a focus on individual performance.

Intragroup Cooperation

Our field sites had not systematically evaluated improvements in performance, but management mentioned instances in which they observed changes in the behavior of workers. For instance, workers were quite willing to suggest solutions to problems encountered in their workgroups, and management sensed that workers' attitudes toward their jobs were better and that tardiness and the number of sick days taken had decreased.

Our discussions with workers and managers reflected the evidence on the benefits of intragroup cooperation in Japanese, Swedish, and American manufacturing organizations (Aguren et al. 1976; Rohlen 1975; Safizadeh 1991). These benefits include better communication and assistance among group members (Schmitt 1984), sharing information to solve problems (Walton 1987), and greater commitment to the organization (Pinto and Pinto 1990; Tjosvold 1984). Although there have been a number of studies on changes in attitudes when firms adopt the intragroup cooperative form of organization, we are not aware of any research on the effects of intragroup cooperation on manufacturing performance. We test this relationship in hypothesis H1:

H1: Cooperative groups will outperform noncooperative groups.

Intergroup Competitive Feedback

Many firms attempt to motivate work groups by having them compete for bonuses or other types of rewards (Doyle 1983; Schmitt 1981; Walton 1987). Our sites were no different. In two of the firms, the rewards were finanical bonuses; in the other firm winning groups received gift certificates. Prior experimental research has shown that intergroup competition can increase performance (Schmitt 1984; Tjosvold 1984).

A key aspect of competition that has not been addressed well at the group level relates to the timing and form of feedback that is received. For individuals, research in social learning theory has shown that the sign (positive or negative) of feedback that people receive is a key determinant of the level of their performance (Bandura 1978; Bandura and Cervone 1983; Ilgen et al. 1979; Podsakoff and Farh 1989). Reactions to performance feedback occur through a self-evaluative mechanism in which comparisons are made between feedback and the performance goals that have been set. Individuals who are committed to explicit performance standards and who obtain negative feedback about performance can become more highly motivated to produce. However, as noted by Bandura and Cervone (1983), performance is not a monotonically increasing function of the discrepancy between negative feedback and performance. If performance is continuously below expectations and individuals come to believe that the standards are not attainable, motivation will decline. But if the discrepancy is more moderate and goals seem attainable, performance can be increased.

In this study we investigate this phenomenon in a group context. Workgroup studies have suggested that competition facilitates performance only when the group believes it has a reasonable chance of winning (Tjosvold 1984). Other researchers have hypothesized that higher performance should result when feedback indicates that the group is either close to the lead or close behind compared to being far ahead or far behind (Schmitt 1984). Hypothesis H2 extends the literature on feedback from the individual to the group level.

H2: Groups who are provided with feedback that they are *just ahead or just behind* will outperform those whose feedback indicates that they are *always ahead or always behind.*

Participation and Slack

Slack has been defined as the difference between an individual's best estimate of performance and the standard chosen when participating in standard selection (Waller 1988; Young 1985;). Since Cyert and March's (1963) conjecture that slack is a reaction to uncertainties in the environment, research has sought to identify other moderating variables that would cause individuals to choose a work standard that was less than their best estimate of performance. For example, an individual's risk preferences, the degree of private information, and the different incentives used cause subjects to choose standards below their best estimates (Kim 1992; Waller 1988; Young 1985).

At all three of our field sites, the reduction of slack was mentioned as a major reason for group competition. Management in particular felt that workers were not performing to their capabilities and were choosing standards that were easily attainable. As a result, they promoted competition by designing incentive schemes that rewarded teams who met their weekly standards. Further, the teams who performed the best over a specified time period (usually a quarter) won a bonus. Feedback was provided to team members regarding the competitions on a weekly basis.

Some of the workers that we talked to told us that the possibility of attaining the bonus had caused them to raise their work standards which is consistent with the literature on feedback. Extending the results of both Bandura and Cervone (1983) and Podsakoff and Farh (1989), we hypothesize that groups that *always* receive feedback that they are leading in a competition will become increasingly satisfied and confident. As more and more positive feedback is received, their self-evaluation and comparison to other groups will cause them to no longer doubt their ability to achieve their goals (Taylor et al. 1984). Thus, they will tend to decrease their effort by building more slack into their standards than those always receiving negative or alternating positive and negative feedback.

H3: Groups receiving feedback that they are *always ahead* of others will build in more slack than those in the other competitive feedback conditions.

II. Experimental Design and Method

Ninety-six undergraduate subjects from advanced cost-accounting classes volunteered to participate in the experiment. Subjects received a fixed number of extra credit points and financial compensation for their participation. The extra credit points were thought to be appropriate as the topic of the experiment coincided with class discussions on management control system design.

Experimental Design

The experiment involves a 3 × 2 completely crossed, repeated measures design. The two independent between-subject variables were: intragroup cooperation (*present, absent*) and intergroup competitive feedback (*always ahead, just ahead or behind,* and *always behind* the leading group). The third factor, time (three periods), was repeated within-subjects.

Intragroup cooperation was operationalized by allowing group members to assist each other through verbal or physical means in performing the task. Subjects in the *noncooperative* groups were not allowed to communicate with or help each other in any manner, they were instructed in performance of only their particular task, and a large wooden divider was set up between subjects so that they could not see their fellow workers. A two-way mirror and a live microphone allowed researchers to directly observe any violations of the cooperation rules, and subjects did not violate these rules.

Intergroup competition was operationalized by having groups vie for a bonus at the end of a series of performance periods. Following each production period, subjects were provided with feedback on how they were doing relative to other groups. The actual performance of groups was not the feedback given to subjects. Although actual feedback would have some advantages, the experimental manipulations used to test the theory could not be done with any degree of control using actual performance.[3]

Competitive feedback was of three types. One group (labeled *always ahead*) always received feedback that they were in the lead; the second group (labeled *just ahead/behind*) was told, in an alternating sequence, that they were either one unit ahead of or one unit behind the leading group; the third group (labeled *always behind*) was always told that they were behind the leading group. The sequence for all three types of feedback was predetermined before the experiment began for control and to allow consistency in the manipulation within each group.[4]

The dependent variables were performance and slack, and they were measured in each of three production periods. Performance was measured as the actual number of good units produced within a period. Slack was measured as the difference between a group's best estimate of their production and the standard they selected.

Experimental Task

The experimental task involved building a toy castle from four subassemblies made from Loc-Blocs.[5] Each of the four subjects in a group was required to build a component and attach it to the subassemblies being passed to him or her. Each subassembly was intricate, and the level of difficulty was set intentionally high (through pretests) so that subjects had to concentrate to produce a final product of "good" quality.[6] Care was taken to ensure that no subject's task was more difficult or time consuming than that of others. Subjects could use any color block to match the required shape. However, to add complexity in construction and to strengthen the cooperation manipulation, one location on each subassembly had to be filled with a block of a particular shape and color. Subjects who were allowed to cooperate could communicate to others or move down the line to ensure that the color and shape matched.

---

[3] Manipulation checks revealed that the feedback stimulated feelings of competition. This procedure has been used for more than 20 years in the psychology literature (see Nadler 1979 for a comprehensive review). The Human Research Committee at the university approved all procedures used in this experiment, and following the experiment we fully informed subjects regarding all manipulations.

[4] Using actual feedback would not enable comparison of the results since each group's performance would probably be different. Thus, there would be no consistent manipulation, and we could not perform statistical tests.

[5] This task has been used in two other experiments (Young 1985; Young et al. 1988).

[6] A completed castle was deemed of good quality if it could pass three tests that subjects were informed of before production. The first involved being able to rotate the castle without it collapsing. In the second, the experimenter made sure that correctly shaped blocks were used in construction. A final test involved inspecting the castle at four key points where prespecified colors and shapes had to be correctly in place. If a castle passed all three of these tests, it was considered a "good product."

Incentive Scheme

The incentive scheme consisted of two components. The first was a payment each period as shown in equation (1). The second component was a bonus that was paid at the end of all production periods. The payment per period was:

$$Y=\$0.10(A)-\$0.10\,|A-S|, \tag{1}$$

where:

Y = the pay for each worker in each group,
S = the group standard chosen by workers, and
A = the actual performance of the group.

Thus, if actual group performance, A, is less than the group standard chosen, S, then compensation is reduced by \$0.10 times the difference between A and S. If A is greater than the group standard, the group makes the same compensation as if it had exactly met its chosen standard. The group can earn more only by raising its standard and meeting it.

This incentive scheme was chosen because it is consistent with the philosophy in many new manufacturing environments in that individuals are encouraged to forecast their production very carefully and then to perform to standard without over- or underproducing (Schonberger 1982). In addition, it has been pretested in previously published research and shown to have strong motivational properties (Young 1985).

To induce intergroup competition, subjects were told that a bonus based on performance relative to other groups would be awarded at the end of the last production period to the group producing the most good quality units.[7] At the end of each period, groups received feedback about their performance relative to others for that period as well as a cumulative performance standing. Since pretests indicated that a subject could earn up to \$12 without the bonus, we set the level of the bonus at \$24 to be split equally among the four team members.[8] The literature on profit and gainsharing indicates that bonuses can range from 5 to 100 percent of base salary in organizations (Doyle 1983); our bonus was set at the midpoint of this range (50 percent of possible base pay) to achieve a high level of motivation and performance.

Taken together, the two types of incentives were designed to (1) promote accurate forecasts of daily performance and impose work discipline (to meet the standard) and (2) generate motivation so that each period's cumulative performance affected a bonus based on sustained performance.

---

[7] As a point of clarification, subjects were told that only the number of good units produced up to the standard they selected would be counted toward the bonus determination. Thus, there was no benefit to producing more than the standard. This is consistent with the daily incentive scheme that discourages production past the standard selected. Teams winning the bonus were determined before the experiment began for control purposes. Thus, actual performance had no effect on the outcome.

[8] We calculated the price paid per unit of production according to the total number of dollars that we had for the experiment and divided this by the number of subjects and periods, allowing for the final bonuses.

Experimental Procedures

Subjects were asked to sign up for a three-day period that included five sessions: a 30-minute training session, one 15-minute trial production period that fully simulated the actual production period, and three measured production periods. We used the trial production period to reduce the effects of differential learning.[9] Subjects were randomly assigned to groups of four and groups were randomly assigned to one of six experimental conditions.[10]

When subjects entered the university's organizational laboratory, an experimenter led them to one of three experimental production rooms. Experimenters were located in adjacent rooms and recorded observations via a two-way mirror. The chief experimenter used headphones to monitor the rooms, and groups were videotaped over the course of the experiment. Before the experimental runs occurred, subjects were informed that they were being observed and that they could be taped. Each experimenter followed a detailed script[11] to maintain consistency across experimenters and to control for researcher expectancy effects (Rosenthal 1966). To avoid subject attrition, we went to great lengths to make sure subjects would return on subsequent days (e.g., we telephoned them each day), and there was no attrition over the three-day period.

Day 1. On the first day, subjects were asked to read a short case (see the Appendix) that discussed the competitive production setting and explained the experimental task. Subjects were instructed as to what they could do to help one another in the cooperation conditions and how they would receive feedback regarding their relative standing. Each subject was randomly assigned to one of four positions on a production line at the beginning of the experiment. Position assignments were permanent; however, subjects in the cooperation condition could move temporarily to help coworkers.

Each subject was given 30 minutes to learn his or her task, and those in the cooperation condition were given extra time to gain familiarity with each other's tasks. Previous research (Young 1985; Young et al. 1988) and pretests indicated that subjects learned their tasks after about 15 minues. After the training session, subjects were asked to produce as many good quality products as they could for 15 minutes; they each earned $0.10 for each good quality castle that the group produced. The flat rate of $0.10 was used to get subjects accustomed to obtaining pay for performance and for motivational purposes. At the end of the period, the experimenter counted and recorded the number of good quality castles produced. The experimenter then told the subjects that they had to wait for the experimenter to tally up the results to see how their group had done relative to other groups. After counting the good units, the chief experimenter returned to the room and told subjects of their standing based on that day's results, and, beginning in the first actual production period, how they were faring cumulatively. As an example, subjects in the *always ahead* group received a relative performance report that stated "Your group is in the lead and X good castles ahead of the next best group. Cumulatively, you are X good castles in the lead." The "Xs" were predetermined each period by the experimenters for each performance report.

---

[9] Waller (1988) observed in his experiment that learning in a three-period experiment decreased after the first period. This bolsters our reasons for eliminating this period from our analysis.

[10] Subjects in the experiment volunteered to participate. Rosenthal and Rosnow (1991) and Birnberg et al. (1990) point out that volunteer subjects may be biased on a particular dimension and not necessarily representative of the general population. Although we did not notice anything unusual about our subjects, there is always the possibility of bias in the results.

[11] This script is 35 pages long and details everything that was said to subjects and all of the actions taken by the experimenters. It is available from the authors on request.

Toward the end of training on the first day, groups were shown the incentive scheme that would be used for the rest of the experiment. The experimenter explained how the scheme worked by going over it several times on a blackboard, then quizzed subjects to ensure that they understood the scheme. The scheme was left on the blackboard so that subjects could refer back to it.

Day 2. On the second day, the trial and first measured production periods, each lasting 15 minutes, occurred. The groups were asked to determine their best estimate of the number of units they could produce. After making this decision, each group wrote this number down and placed the information in an envelope and sealed it. At this point, in conformance with the ideas that competitive feedback would mediate the relationship between the best estimate and the actual standard selected, the experimenter reentered the room and performed the competitive feedback manipulation by revealing the group's performance standing based on the previous day's production. Given the relative standing information, group members then negotiated with each other until a unanimous decision about their work standard was reached. The group then informed the experimenter of the group standard, which was recorded. Since slack was defined as the best estimate of performance minus the actual standard selected, these procedures provided us with the measure of group slack.

Subjects then worked for a 15-minute period, after which a questionnaire was administered to gather information relating to other behavioral variables of interest.[12] At the end of the second period on Day 2, the competitive feedback procedures were followed.

Day 3. The second and third measured production periods on Day 3 used the same procedures as outlined for the first measured period. To avoid any end-game strategies or motivational problems, subjects were not told for how many periods the game would continue. At the end of period 4, the (predetermined) group winning the weekly bonus was announced, and subjects then filled out a final questionnaire, which contained the same questions as the previous questionnaires plus manipulation checks. Finally, subjects were thoroughly informed about the purposes of the experiment and the manipulations that were used, and each subject who did not receive the bonus was given an additional $4 to equalize payments across subjects.[13]

III. Results

Even though we randomly assigned subjects to experimental conditions, we performed analyses of variance on demograhic data related to subjects. This data included age, gender, academic major, amount of full-time and part-time work experience, and the number of accounting courses subjects had taken. The results indicated no significant differences across cells, and thus no systematic bias in the experimental results.

---

[12] As this is one of the first studies in this area, we decided to use the questionnaires between periods to develop behavioral constructs related to group cohesion and commitment to standards for future research.

[13] Across all experimental conditions, subjects made $11.25 on average.

Manipulation Checks

The manipulation of competitive feedback was assessed by subjects' responses to two statements, the first of which was "It seemed that our team was always leading other groups." Subjects responded on a seven-point Likert scale that ranged from "strongly disagree" (1) to "strongly agree" (7). An ANOVA was used to test for significant differences across conditions. A main effect was found for competition ($F = 124.99$, $p < 0.000$). The marginal means for the three competitive manipulations across cooperative conditions were in the right direction: *always ahead*, $\bar{x} = 6.77$; *just ahead/behind* $\bar{x} = 4.60$, and *always behind* $\bar{x} = 1.78$. Thus, the feedback element of the competition manipulation was shown to work as expected.[14]

The second statement, "The performance of other teams influenced our team's production goals," was designed to assess whether standard setting was affected by competitive feedback. The same scale was used as before. An ANOVA across conditions indicated a main effect for competitive feedback ($F = 6.768$; $p < 0.002$). Again, the marginal means were in the right direction: *always ahead*, $\bar{x} = 4.47$; *just ahead/behind*, $\bar{x} = 5.63$; *always behind*; $\bar{x} = 5.72$. Contrasts among the groups indicated significant differences between the *always ahead* and *just ahead/behind* groups ($t = 3.303$, $p < 0.001$) and between the *always ahead* and *always behind* groups ($t = 3.05$, $p < 0.003$). No difference was found between the *just ahead/behind* and *always behind* groups ($t = 0.248$, $p < 0.800$). It is clear from this manipulation check that those *just ahead/behind* and *always behind* were more strongly affected by the feedback compared to those in the *always ahead* condition. Table 1 contains descriptive statistics.

The cooperation manipulation check used four statements combined into one scale. These were: "No one is allowed to help me if I run into trouble on my task during production" (negatively scored); "I can help others in performing their tasks during production;" "I can see how well my other team members are performing during production;" and "I can communicate freely with others during production." Subjects responded to each statement on a seven-point Likert scale that again ranged from "strongly disagree" (1) to "strongly agree" (7). Scale reliability for this four-item scale was quite satisfactory (Cronbach's alpha $= 0.82$).

The four items in the scale were summed and the resulting variable was used as the dependent measure in an ANOVA across the experimental conditions. The results show a significant effect for cooperation ($F = 68.14$, $p < 0.0001$). The marginal mean for groups in the cooperation condition was 19.30, while that for noncooperation, 18.83, indicates that the significant difference was in the desired direction. These results indicate a successful manipulation of cooperation.

Table 1 presents descriptive statistics for the best estimate of performance, actual group standard selected, slack, and actual performance. Several patterns are worth nothing. First, best estimates and group standards increased over time in almost all of the experimental conditions. Second, the range of best estimates and standards is higher for the *just ahead/behind* and *always behind* groups than for those *always ahead*. Third, when comparing group best estimate changes (period 4 vs. period 2), those given *always ahead* feedback had a smaller change than the other two groups. The pattern of

---

[14] We expected subjects who were *just ahead/behind* to feel that they were sometimes in the lead and sometimes not. This would cause them to fall between the *always ahead* and *always behind* groups.

Table 1
Descriptive Statistics for the Six Experimental Conditions
(n = four groups per cell with four subjects per group)

	Cooperative Feedback Manipulation					
	Cooperative Groups			*Noncooperative Groups*		
	Ahead	*Ahead/ Behind*	*Behind*	*Ahead*	*Ahead/ Behind*	*Behind*
Group Best Estimate:						
Period 2						
mean	16.75	18.50	16.50	15.75	17.75	18.25
s.d.	5.13	2.37	2.13	2.35	4.28	2.34
Period 3						
mean	17.25	22.25	19.75	16.50	19.75	20.00
s.d.	4.09	1.84	2.46	1.71	4.47	3.10
Period 4						
mean	17.50	22.00	21.75	16.25	22.00	20.00
s.d.	3.31	2.53	1.13	1.53	4.90	1.47
Group Standard:						
Period 2						
mean	16.50	18.50	16.75	14.75	17.75	18.75
s.d.	4.76	2.37	2.11	1.98	4.22	0.86
Period 3						
mean	17.00	19.75	20.75	15.75	19.75	19.75
s.d.	4.44	1.13	2.67	1.53	4.22	2.96
Period 4						
mean	16.00	20.75	22.25	16.25	22.00	19.50
s.d.	3.80	0.86	0.45	1.84	4.84	1.71
Group Slack:						
Period 2						
mean	0.25	0.00	−0.25	1.00	0.00	−0.50
s.d.	0.45	0.00	0.45	1.79	0.73	1.55
Period 3						
mean	0.25	2.50	−1.00	0.75	0.00	0.25
s.d.	0.45	1.55	1.63	1.34	0.73	0.86
Period 4						
mean	1.50	1.25	−0.5	0.00	0.00	0.50
s.d.	2.13	1.98	1.16	0.73	0.73	0.89
Actual Production:						
Period 2						
mean	16.75	17.50	19.00	17.75	18.75	18.75
s.d.	3.45	1.55	1.93	2.77	5.18	3.45
Period 3						
mean	16.25	19.25	17.00	15.50	20.50	19.00
s.d.	3.04	1.13	1.27	0.52	5.34	1.93
Period 4						
mean	15.50	21.25	19.50	17.50	22.75	21.00
s.d.	2.69	0.86	2.78	1.55	5.53	1.46

change in group standards is consistent with the results of the best estimates. Those groups given *always ahead* feedback tend to have lower group standards, and these standards do not change as substantially in comparison with the other groups.

Table 2
MANOVA Results for the Effects of Cooperation and Competition on Actual Production and Slack

	Actual Production		Slack	
Factor	*F-Statistic*	*Probability*	*F-Statistic*	*Probability*
Between Subjects				
Cooperation	3.91	0.052	1.75	0.189
Competition	14.91	0.000***	12.07	0.001***
Cooperation by Competition	0.10	0.913	11.03	0.001***
Within Subjects				
Cooperation by Time	3.40	0.038*	2.60	0.082
Competition by Time	12.30	0.001***	5.80	0.000***
Cooperation by Competition by Time	3.77	0.006**	3.70	0.006**
Time	45.24	0.001***	4.59	0.013*

* Significant at the 0.05 level.
** Significant at the 0.01 level.
*** Significant at the 0.001 level.

Patterns in the data for group slack are more difficult to detect. Regarding actual production, the *always ahead* group produced consistently less than the other two groups and its production in period 4 versus period 2 actually dropped.

Tests of Hypotheses

Hypothesis H1: Hypothesis H1 predicts that subjects in cooperative conditions will perform better than those in noncoopertive conditions. Results of the repeated measures MANOVA (with time as the repeated measure) showed a marginally significant main effect for cooperation ($F=3.91$, $p<0.052$; see table 2); however, no subjects in the noncooperative groups produced more than those in the cooperative condition.[15] An explanation for this may be that subjects in the noncooperative condition had more time to produce output since they were not allowed to move to help one another. Workers who were cooperating had to physically move from one station to the next, which seemed to slow production, and in some cases they stopped to discuss their performance. But those who were unable to cooperate stayed in the same location and simply produced more output. Although this result seems surprising, given anecdotal evidence from industry, the results suggest that cooperation may be effective only in certain situations. For example, if the production line is not well-balanced, then cooperation may aid production managers in finding bottlenecks as more individuals are needed in certain parts of the line than in others. Cooperation may also lead to improvements once a line is balanced by reducing the number of workers required to perform a particular task. These and other possibilities can be explored in future studies.

---

[15] In testing hypotheses H1 and H3 with repeated measures MANOVA, Mauchly's sphericity test revealed that the assumption of symmetric covariance matrices (equal variances on the diagonal and zero covariances off the diagonal) was violated. As a result, we were forced to use the multivariate as opposed to univariate tests to assess effects (see Latour and Miniard 1983 for a discussion).

Table 3
Linear Contrasts by Period

Contrast	*Period 2*	*Period 3*	*Period 4*
Actual Production			
Ahead (A) vs. Just Ahead/Behind (AB)	$t=1.067$ $p<0.289$ N.S.	$t=5.86$ $p<0.000$ A/B>A	$t=7.56$ $p<0.000$ A/B>A
Ahead (A) vs. Behind (B)	$t=1.982$ $p<0.05$ B>A	$t=3.115$ $p<0.002$ B>A	$t=5.155$ $p<0.000$ B>A
Just Ahead/Behind (AB) vs. Behind (B)	$t=0.915$ $p<0.363$ N.S.	$t=2.749$ $p<0.007$ A/B>B	$t=2.41$ $p<0.018$ A/B>B
Budgetary Slack			
Ahead (A) vs. Just Ahead/Behind (AB)	$t=2.396$ $p<0.019$ A>A/B	$t=2.55$ $p<0.013$ A/B>A	$t=0.359$ $p<0.720$ N.S.
Ahead (A) vs. Behind (B)	$t=3.83$ $p<0.000$ A>B	$t=2.97$ $p<0.004$ A>B	$t=2.15$ $p<0.034$ A>B
Just Ahead/Behind (AB) vs. Behind (B)	$t=1.44$ $p<0.154$ N.S.	$t=5.52$ $p<0.000$ A/B>B	$t=1.79$ $p<0.076$ A/B>B

Hypothesis H2: This hypothesis states that workers who receive feedback that they are *just ahead or behind* the leading group will outperform those in the other competitive feedback conditions. To be most informative, we report the results of this hypothesis on a period-by-period basis rather than by summing across periods. In addition, since competitive feedback interacts with time, it is inappropriate to simply sum the performance variables across the three periods and perform a single linear contrast. We performed a series of linear contrasts comparing the *just ahead/behind* group to the *always ahead* and *always behind* groups, respectively, for all three performance periods. Table 3 shows the results of all the contrasts by period. In period 2, the *just ahead/behind* group did not produce significantly more than the *behind* group ($t=0.92$, $p<0.05$), nor did they produce significantly more than the *always ahead* group, although the marginal means show that the difference was in the predicted direction ($\bar{x}=18.175$ for the *just ahead/behind* group, $\bar{x}=17.25$ for the *always ahead* group). Thus, the hypothesis was not supported for period 2. However, for both periods 3 and 4, the contrasts show strong support for the hypothesis, with the *just ahead/behind* group significantly outperforming both the *always ahead* group and the *always behind* group.

Hypothesis H3. Hypothesis H3 suggests that workers receiving competitive feedback that they are *always ahead* of the next best group will build in more slack than those in the other competitive conditions. In table 2 the results of the repeated measure MANOVA show that both the competitive feedback main effect and interactions are significant ($p < 0.01$). The results of the contrasts by period are shown in table 3. For the second period, the hypothesis holds, with the *always ahead* group producing more slack than either of the other groups. In period 3, the hypothesis is only partially supported because, although the *always ahead* group produced more slack than the *always behind* group, the *just ahead/behind* group created significantly more slack than the *always ahead* group. In the fourth period, the *always ahead* group created more slack than the *always behind* group, but no more than the *just ahead/behind* group. Overall, there is moderate support for this hypothesis.[16]

IV. Discussion

Using insights gained from site visits to major manufacturing organizations and the current literature on manufacturing, we have examined how cooperation and competitive feedback affect workgroup performance and slack. As noted by Safizadeh (1991), research on workgroups has been difficult to interpret because of lack of control in some field research and/or poor experimental designs. Because we performed this study in the laboratory, we were better able to manipulate the variables of interest while holding other variables constant. We believe that this study does have implications for firms employing workgroups, as discussed below.

The intragroup cooperation hypothesis was not supported, and the results were in the opposite direction to that predicted. The results suggest that the form of cooperation that we manipulated may not yield greater output than noncooperation on this task. While this is surprising, we observed from behind the two-way mirrors in the laboratory that cooperating subjects did spend time moving from their work stations to help each other. Further, there was a certain amount of morale boosting, through verbal encouragement, that occurred in these groups. In contrast, noncooperating groups did little else but work hard to meet their standard. These subjects appeared very intense and the mood in these groups was quite somber. It could be that if the experiment had lasted significantly longer, the cooperative groups would have outperformed the noncooperative groups.

Our results indicate that the cooperation construct needs to be investigated more thoroughly before clear conclusions about its effectiveness can be drawn. It seems that cooperation may be effective only under certain conditions and for specific kinds of tasks. Although we were careful to select a difficult task and attempted to "balance the line" by performing pretests, it appears from Chase and Aquilano (1988) that one major benefit of cooperation is to aid manufacturing designers over time in developing a balanced line. It could be that cooperation improves group performance only when the line is unbalanced and some individuals have more difficult tasks to perform than others do. However, since production management texts indicate that line balancing is a central part of a good manufacturing process (Chase and Aquilano 1988) and that production lines at major manufacturing facilities such as Toyota are balanced (Schonberger 1982), we felt that cooperative effects would continue to be derived even with the line that we used.

---

[16] We should note that the best estimates of the *always ahead* group tend to be lower than those of the other two groups. In fact, in the last production period, the *always ahead* group is lower than the other two groups. It is possible that lower performance in previous periods affected the group's perception of their best estimate. This possibility does not affect our definition of slack as a conscious decision to set the standard below perceived best performance. However, this result does hint at a potentially interesting interaction between prior performance and the perceived best estimate which could be studied in a future experiment with a larger number of performance periods. We thank one of the reviewers for suggesting this.

The hypotheses regarding competitive feedback were for the most part supported. Those groups receiving *just ahead/behind* feedback tended to produce more units than the other groups. Those groups receiving feedback that they were *always behind* also produced more units in all three periods than the groups receiving feedback that they were *always ahead*. The results regarding competitive feedback extend the work of social learning theorists (e.g., Bandura and Cervone 1983; Podsakoff and Farh 1989) to a multi-period group context and provide preliminary evidence on the strong motivation associated with alternating the sign of feedback.

An implication for organizations employing internal competition is that some level of competition is good for motivation, but workgroups that begin to fall behind may become debilitated over time and production may decrease. Conversely, some workgroups may excell in performance. This imbalance may cause coordination problems both within a division and across divisions. Thus, organizations will have to consider how competitions affect overall manufacturing performance and strategic and production planning.

Further, it should be noted that although significant differences in budgetary slack were found across some experimental conditions, the relative amount of slack built into standards was quite low (see table 1). It appears from these results that the low levels of slack occurred because workgroups were highly motivated to attain both a high level of compensation based on their daily standards and the final bonus. If firms desire to reduce slack, our findings are that offering the right combination of incentives to workers will provide this result.

Finally, as suggested in many studies on the effects of incentives, there is the possibility that individuals with a high need for achievement may at some point no longer be motivated by money; more pay may not result in more work (McClelland 1972, 1985). Such individuals are affected by financial rewards but tend to use them as a benchmark for social comparison with others. Social factors such as prestige and position will often override financial factors. Although such issues related to individual characteristics can be overcome via random assignment to experimental conditions, future research should consider measuring individual need for achievement to understand more clearly how this important variable affects the results of research on the motivational effects of incentives.

Appendix

TOYCO Corporation

Castle Division Case

Castle Division is the newest division of the TOYCO Corporation. It produces a new product line of toy castles. The toy castles that are produced in this division are part of a set of toys, including toy bridges, knights and dragons, all of which have a medieval theme. These other toys are made in different divisions. Production of the castles has to be carefully planned so that the total number of castles will match the total number of other toys (one castle per set).

Castle Division was established in November. Because the toy castle is the newest entry into the medieval theme market, management is still not sure how many *good quality* castles it can produce. Management needs to have this information in order to determine how many orders it can fill for Christmas.

Castle Division has selected you to be a production trainee. Your job is to work with three other employees in a team to make castles. Remember that each subassembly of the castle that you make, and each final product, has to be of the *highest quality*. A poor quality castle *cannot* be counted in production and will ultimately decrease your compensation for the job. At the end of this training period, only the teams with the best performance will be selected as full-time employees. You will be competing against four or five other teams.

In addition to their daily compensation, the best performing teams will receive an end-of-training financial bonus. Management is seeking employees who can be hired and trained so that full production can begin in mid-December. Remember, it is your team and not you as an individual that is eligible to be hired as full-time employees. You will be instructed in how to make such castles shortly.

References

Aguren, S., R. Hansson, and K. G. Karlsson. 1976. *The Impact of New Design on Work Organization*. Stockholm, Sweden: The Rationalization Council SAF-LO.

Bandura, A. 1978. The self system in reciprocal determinism. *American Psychologist* 33 (June): 344–58.

——, and D. Cervone. 1983. Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology* 45 (September): 1017–28.

Birnberg, J. G., M. D Shields, and S. M. Young. 1990. The case for multiple methods in empirical management accounting research (with an illustration from budget setting). *Journal of Management Accounting Research* 2: 33–66.

Chase, R., and N. Aquilano. 1988. *Production and Operations Management—A Life Cycle Approach*. Homewood, IL: Richard D. Irwin.

Chow, C. W. 1983. The effects of job standard tightness and compensation scheme on performance: An exploration of linkages. *The Accounting Review* 58 (October): 667–85.

——, J. C. Cooper, and W . S. Waller. 1988. Participative budgeting: Effects of a truth inducing pay scheme and information asymmetry on slack and performance. *The Accounting Review* 60 (January): 111–22.

——, ——, and K. Haddad. 1991. The effects of pay schemes and ratchets on budgetary slack and performance: A multiperiod experiment. *Accounting, Organizations and Society* 15 (January): 47–60.

Cyert, R., and J. March. 1963. *The Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.

Deutsch, M. 1949. The effects of cooperation and competition upon group processes. *Human Relations* 2 (January): 129–52.

Doyle, R. J. 1983. *Gainsharing and Productivity*. New York: AMACOM.

Goldman, M., J. W. Stockbauer, and T. G. McAuliffe. 1977. Intergroup and intragroup competition and cooperation. *Journal of Experimental Social Psychology* 20 (January): 81–88.

Hayes, R. H., S. C. Wheelwright, and K. B. Clark. 1988. *Dynamic Manufacturing*. New York: Free Press.

Hoerr, J. 1989. The payoff from teamwork. *Business Week* (July): 56–62.

Ilgen, D. R., C. D. Fisher, and M. S. Taylor. 1979. Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology* 64 (March): 349–71.

Johnson, D. W., G. Maruyama, G. R. Johnson, D. Nelson, and L. Skon. 1981. Effects of cooperative, competitive and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin* 90 (January): 47–62.

Kim, D. C. 1992. Risk preferences in participative budgeting. *The Accounting Review* 67 (April): 303–18.

Latour, S. A., and P. W. Miniard. 1983. The misuse of repeated measures analysis in marketing research. *Journal of Marketing Research* 20 (January): 45–57.

McClelland, D. C. 1972. The role of money in managing motivation. In *Managerial Motivation and Compensation*, edited by H. Tosi et al. East Lansing, MI: Michigan State University Business Studies.

———. 1985. *Human Motivation*. Glenview, IL: Scott Foresman.

Miller, L. K., and R. L. Hamblin. 1963. Interdependence, differential rewarding and productivity. *American Sociological Review* 28 (October): 768–78.

Nadler, D. A. 1979. The effects of feedback on task group behavior: A review of the experimental research. *Organizational Behavior and Human Performance* 31 (July): 309–38.

Okun, M. A., and F. J. DiVesta. 1975. Cooperation and competition in coacting groups. *Journal of Personality and Social Psychology* 32 (October): 615–20.

Pinto, M. B., and J. K. Pinto. 1990. Project team communication and cross-functional cooperation in new program development. *Journal of Innovative Management* 5 (March): 200–12.

Podsakoff, P. M., and J. Farh. 1989. Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes* 24 (January): 45–67.

Rohlen, T. 1975. The company work group. In *Japanese Organization and Decision Making*, edited by E. Vogel. Berkeley, CA: University of California Press.

Rosenthal, R. 1966. *Experimenter Effects in Behavioral Research*. New York: Appleton-Century-Crofts.

———, and R. L. Rosnow. 1991. *Essential of Behavioral Research: Methods and Data Analysis*. New York: McGraw-Hill.

Safizadeh, M. H. 1991. The case of workgroups in manufacturing operations. *California Management Review* 33 (Summer): 61–82.

Schmitt, D. R. 1981. Performance under cooperation and competition. *American Behavioral Scientist* 22 (November): 649–79.

———. 1984. Interpersonal relations: Cooperation and competition. *Journal of the Experimental Analysis of Behavior* 10 (September): 377–83.

Schonberger, R. 1982. *Japanese Manufacturing Techniques—Nine Hidden Lessons in Simplicity*. New York: Free Press.

———. 1986. *World Class Manufacturing*. New York: Free Press.

Steiner, I. D. 1972. *Group Process and Productivity*. New York: Academic Press.

Taylor, M. S., C. D. Fisher, and D. R. Taylor. 1984. Individuals' reactions to performance feedback in organizations: A control theory perspective. In *Research In Personnel and Human Resources Management*, edited by K. M. Rowland and G. R. Ferris. Greenwich, CT: JAI Press, 81–124.

Tjosvold, D. 1984. Cooperation theory and organizations. *Human Relations* 37 (December): 743–67.

Waller, W. S. 1988. Slack in participative budgeting: The joint effects of a truth-inducing pay scheme and risk preferences. *Accounting, Organizations and Society* 13 (January): 87–98.

———, and C. W. Chow. 1985. The self-selection and effort effects of standard-based employment contracts: A framework and some empirical evidence. *The Accounting Review* 60 (July): 458–76.

Walton, R. E. 1987. *Innovating to Compete*. San Francisco: Jossey-Bass.

Young, S. M. 1985. Participative budgeting: The effects of risk aversion and asymmetric information on budgetary slack. *Journal of Accounting Research* 23 (Autumn): 829–42.

———, M. D. Shields, and G. Wolf. 1988. Manufacturing controls and performance: An experiment. *Accounting, Organizations and Society* 13 (October): 607–18.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 482–514

Experimental Evidence on Tax Incentives and the Demand for Capital Investments

Jon S. Davis
University of Colorado at Boulder
Charles W. Swenson
University of Southern California

SYNOPSIS AND INTRODUCTION: Since the pioneering work of Hall and Jorgenson (1967), numerous studies (e.g., Bischoff 1971; Chirinko and Eisner 1982; and Coen 1971) have examined the effect of attempts by the tax authority to influence investment decisions through accelerated depreciation or investment tax credits (ITC). This body of research has been fraught with econometric estimation problems, and consequently has failed to provide a clear picture of the effect of tax policies on capital investment. In a review of the literature, Chirinko (1986, 151) concludes that "[w]hile investment may respond significantly to variations in tax parameters, it appears to this author that the supporting empirical evidence has yet to be generated."

At the core of the difficulties in the econometric research paradigm is the operationalization of the neoclassical investment function itself. Chirinko (1986) notes that numerous inherent difficulties are introduced, including (1) estimations of the purchase cost of a unit of capital, financial cost of capital net of inflation, rate of depreciation of the capital good, rate of income taxation, rate of investment credit, discounted value of depreciation allowances, net cost of debt finance, and the like; and (2) the inability to control for firms' expectations regarding output, and hence the marginal product of capital. These difficulties highlight the general limitations of econometrics in certain settings. This sentiment was echoed by Chirinko and Eisner (1983, 139) when they concluded that, in the neoclassical tax

Comments welcome. The authors acknowledge the aid of Herman Quirmbach and the helpful comments of workshop participants at University of Arizona, Arizona State University, and San Diego State University. The authors also acknowledge the generous financial support of the University of Southern California, Investors in Business Education at the University of Illinois, and especially the KPMG Peat Marwick Tax Research Opportunities Program. In addition the research assistance of Minou Bohlin is gratefully acknowledged.

Submitted December 1991.
Accepted January 1993.

policy arena, "one can get almost any answer one wants by making sure that the chosen model has specifications appropriate to one's purpose."

In response to the inconclusive econometric evidence regarding the effect of tax incentives on capital investment, we adopt an alternative approach in this study, using laboratory markets to overcome the limitations noted above, thereby providing a controlled empirical test of neoclassical predictions. Although the results of our experiments provide no evidence regarding the real-world dollar responses of investment to income tax accounting subsidies, some insight into the ability of theory to predict more general aspects of taxpayer investment behavior is provided. Specifically, the research question addressed is whether capital investment increases when depreciation or investment credits allowed by the tax system result in more rapid deductions than true economic depreciation. Although this question follows directly from neoclassical predictions, we relax the assumption of price taking to permit the more realistic consideration of market price adjustments.

The results of our experiments do not support the neoclassical prediction that depreciable asset investment will increase in response to accelerated tax depreciation or to investment tax credits. Demand was unresponsive to tax incentives because the prices of depreciable assets were bid up. That is, tax benefits were captured to some extent by factor suppliers. From a theoretical perspective, the study's results provide a "piece of the puzzle" in light of conflicting or nonexistent econometric evidence.

In section I, a description of the experimental setting and administration is provided. Theoretical predictions of investment price and quantity are then derived from our experimental operationalization of a production economy in section II. Finally, results and conclusions are presented in sections III and IV, respectively.

Key Words: *Capital investment, Experimental economics, Depreciation, Tax.*

Data Availability: *The data upon which this article is based may be obtained from the authors on request.*

I. The Experimental Setting

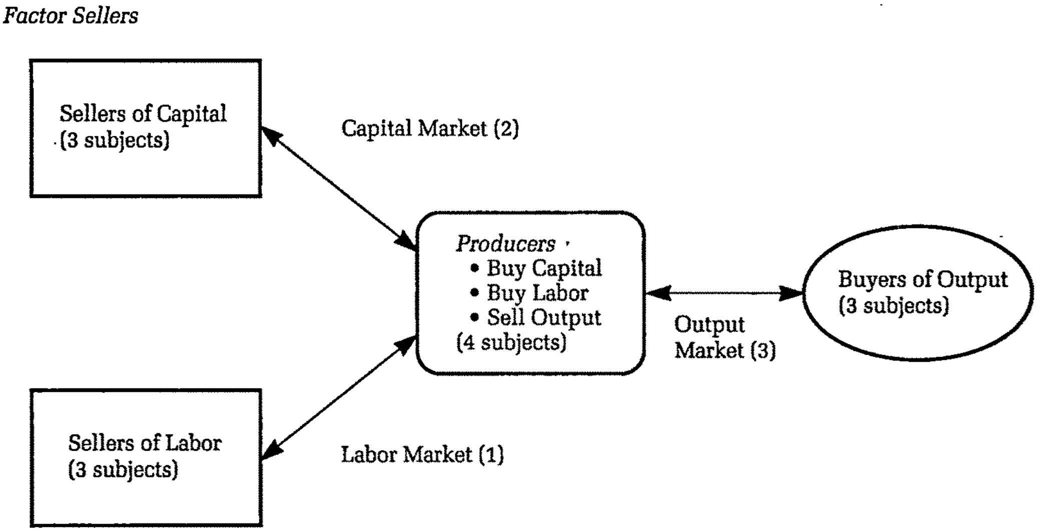
The Environment

SIX separate experimental sessions were utilized, each featuring a production economy with three classes of economic agents: factor sellers, output producers, and output buyers. Six factor sellers were initially endowed with factor inputs three sellers each of capital and labor) that were sold to four producers who were nitially endowed with cash. Producers combined their purchases of capital and labor to produce output that was subsequently sold to three buyers, also initially endowed with cash. These transactions occurred in three separate multiple-unit, double-auction markets.[1] The markets are represented graphically in figure 1. At the start of a market

---

[1] The double-auction institution was chosen to facilitate a stable competitive equilibrium price. See Davis and Swenson (1988, 46) for a description of this institution.

Figure 1
The Experimental Production Economy



period, the capital and labor markets began simultaneously and operated until producers indicated that they no longer wished to purchase additional factor inputs. During the operation of the two factor markets, producers' endowments of output inventory were increased in discrete units as capital and labor were purchased, according to the following (decreasing-returns-to-scale) Cobb-Douglas production function:

$$F(x_1,x_2)=A(x_1,x_2)^{\zeta}, \tag{1}$$

where F denotes ouput, and x_1 and x_2 are labor and capital inputs, respectively, A is a scale factor (set to 2.27), and ζ is an elasticity coefficient (set equal to 0.4).[2] Immediately after the factor markets were closed, a separate output market was opened in which producers' inventory was sold to output buyers.[3]

Following the close of the output market, producers sold their used capital to the experimenter at a prespecified salvage value, output buyers resold their inventory to the experimenter at a prespecified price, and all subjects computed their profits, a fixed proportion of which was paid in cash at the end of the experiment. Profit for capital and labor sellers was computed by the application of the following two equations, respectively:

---

[2] The decreasing-returns-to-scale form of the production function (i.e., the sum of capital and labor coefficients set to less than 1) was chosen to induce an increasing marginal cost function, which in turn allows for a unique solution for producers (i.e., a single intersection of the marginal cost and output demand curves).

[3] The nonsynchronous input and output markets were necessary because the software used could not calculate continuous marginal product with the above production function. See the conclusion for additional discussion on this point.

$$\pi f, x_2 = \sum w_{2j} - \sum \kappa_j, \tag{2a}$$

$$\pi f, x_1 = \sum w_{1m} - \sum \lambda_m, \tag{2b}$$

where π_f is the profit for factor sellers, w_{2j} is the sales price of unit j of new capital, w_{1m} is the sales price of unit m of labor, κ_j is the initial cost to capital sellers for unit j of new capital (set exogenously by the experimenters), and λ_m is the initial cost for unit of labor for labor sellers (set exogenously by the experimenters). Thus, the remuneration paid to factor sellers was based on the difference between sales proceeds of labor or capital and the initial cost for these factors. Factor inventory that was not sold during the market period had zero value at the conclusion of the period and could not be carried forward to future periods. Profit for output buyers was determined by subtracting the cost of purchased units from the sales proceeds received when their purchases were resold to the experimenter at exogenously determined prices (redemption values).

Profit for producers, π_p, was determined by the following equation:

$$\pi_p = \sum p_i - \sum w_{1i} - \sum (1-\beta) w_{2i}, \tag{3}$$

where p_i is the sales price of one unit of output, and β is a bonus (defined below). Thus, producers' profits are revenues from the sale of output, less costs paid for labor and capital. Capital cost is net of a "bonus" as a percentage of total capital purchases. The bonus is composed of two terms: an economic salvage value (set at 25 percent of the purchase price of capital) and a tax benefit, set at either zero (representing either no ITC or tax depreciation equal to economic depreciation) or 25 percent of the purchase price (representing the tax benefit generated by accelerated depreciation or ITC). As with unsold factor inventory, output inventory that was not sold during the market period had zero value at the conclusion of the period and could not be carried over to future periods.

Our characterization of tax incentives are equivalent to either ITC or depreciation when using a single-period model, since the asset is fully depreciated and sold at the end of each period. We note that in the real world, any difference in tax depreciation treatment is eventually repaid in its entirety in the form of a taxable gain on sale. The tax benefits thus follow from the time value of money; the longer the asset is held, the lower the present value of the tax on the gain. With respect to depreciation, the assumption in this experiment is that a number of years are compressed into one. We also simplify producers' (as well as factor sellers' and output buyers') computations by not having them reduce their overall profits by a tax rate. It is well-known that a flat tax rate on profits *per se* has no theoretical effect on producers' output (or input decisions), sellers' sales volume, buyers' purchases, or any derivative market prices.

Factor sellers' initial technological endowment consisted of private knowledge of their individual cost schedules for capital and labor. Producers' technological endowment consisted of private knowledge of their production functions. Output buyers' technological endowment consisted of private knowledge regarding the resale price schedules for output.

Experimental Administration and Design

The experiments used 78 volunteers enrolled in upper-division undergraduate business school courses at a major metropolitan university. Subjects with previous double-auction experimental market experience were assigned the role of producers because of the complexity of the task. All other subjects had no previous laboratory market experience, and were randomly assigned buyer or seller roles.

The site of the experiment was an experimental economics laboratory with a system of networked personal computers. Because of the complexity of the task, subjects were required to participate in a 90-minute training session the day before the experiment. After reading written instructions, subjects completed an instructional program that used interactive examples to explain the operation of the markets, the subject's role and task in the experiment, and how profits and cash payments would be determined. Buyers and sellers also participated in a simulated market, which was played against each individual subject's computer. While other subjects completed the simulated market training, the producers were instructed on the use of the production function on the computer, and engaged in buying, producing, and selling exercises with one of the experimental proctors.

After completing the instructions, subjects engaged in two practice markets in a production economy. They used a modified version of the networked Multiple Unit Double Auction (MUDA) program developed at the California Institute of Technology (Plott 1991). MUDA allows for nearly instantaneous buying and selling of assets in multiple markets. The program displays sales prices, acquisition costs, and all bids and offers from all buyers, sellers, and traders on an ongoing, historical basis. In addition, for each subject, both the current standing bid and offer (which were determined using a convergence rule) in each market and cash and inventory balances are displayed.[4] To avoid real-world referents in the instructions (to discourage role playing by subjects and maintain control over preferences), terms such as "units from markets 1 and 2" and "bonus" were used in lieu of "capital," "labor," "tax benefit," and the like.

On the day following training, subjects returned for the actual experiment. Time to complete the experiment was three hours. All subjects were paid in cash, according to their profits, at the end of the experiment. Average compensation for the entire experiment was $72.21 per subject, with a low payment of $30 and a high payment of $200.[5] Because each market period was nine minutes long, there was sufficient time for only 12 to 14 trading periods in each experiment.

Induced supply conditions were assigned to factor input markets by assigning costs to factor units on the basis of the following isoelastic supply functions:

$$x_1^s = a \cdot w_1, \qquad \text{for labor,} \tag{4}$$

---

[4] The MUDA program (and a number of supporting programs and documentation) are available upon request from Charles Plott, Social Sciences Division, California Institute of Technology, Pasadena.

[5] Although $72 per subject may appear to be excessive for participation in the experiment, it represented slightly less than $15 per hour (including training time), which is similar to the compensation in other experimental economics studies. In addition, the experimenters felt that $15 per hour was necessary to ensure control over subject preferences, given the complexity of the markets and length of time required. The conversion rates for the experimental currency were as follows: producers, one franc to $0.12; output buyers, one franc to $0.005; Market 1 and 2 sellers, one franc to $0.03. The range of dollar payments to subjects, by group, was: Producers—$30 to $180; Market 1 sellers—$31 to $125; Market 2 sellers—$35 to $133; and Market 3 buyers—$56 to $200. A post-experimental questionnaire suggested that all of Smith's (1982) experimental precepts were satisfied.

and

$$x_2^s = b \cdot w_2, \qquad \text{for capital,} \tag{5}$$

where x^s represents the discrete number of goods in inventory, $a=5.12$, $b=0.72$, and induced cost, w, is derived accordingly. Each of the three sellers of labor (capital) was endowed with 15 (ten) units of inventory, more than enough to satisfy predicted demand.

Induced demand for the output buyers was of the following isoelastic form:

$$y^D = \frac{\alpha}{p}, \tag{6}$$

where α was set to 500 and y was in discrete units. Redemption values, p, were then derived from this function. Each of the three output buyers was endowed with 3,000 francs (experimental currency), more than enough to purchase all output produced at the theoretically predicted level.[6]

Since the hypotheses focus on the effect of accelerated tax depreciation, two experimental conditions were created. First, three experimental sessions were run either with no tax benefit provided or with tax and economic depreciation both set at 75 percent of cost. Thus, the bonus (β) was set at 25 percent of cost to represent salvage value. In three other sessions, the additional effect of ITC or accelerated tax depreciation was added to the bonus, which resulted in a bonus equal to 50 percent of cost.[7]

II. Theoretical Predictions

Continuous-Form Solutions

To determine the effects of the introduction of a tax benefit, we first determine competitive equilibrium (CE) prices and quantities in all three markets without tax benefits (i.e., tax and economic depreciation equality). These CE prices and quantities are then compared to those predicted under the tax benefit condition. In either condition, this process implies solving for the points at which all markets' supply and demand curves simultaneously cross. As noted, schedules for factor market costs (supply) and output redemption (demand) were provided exogenously, which is similar to the induced supply and demand schedules commonly used in single-market experiments (see Smith 1982).

For producers, no such schedules for redemption (for the input market) or cost (for the output market) were provided. Instead, induced demand and supply is determined endogenously. In the following discussion, we derive continuous-form solutions for producers and equate them to the supply and demand schedules provided to factor sellers and output buyers.

---

[6] Because producers would not initially know what prices would be in the output markets, there might be some tendency to over- or underpurchase from the input markets (and over- or underproduce). To encourage producers to explore various input-output combinations, they were told that, for the first two periods only, any losses would be treated as zero profit. Beyond the second period, however, any losses incurred for a period strictly offset profit in the other periods.

[7] A bonus was used rather than requiring subjects to compute depreciation and tax liabilities because (1) given the complex and time-consuming nature of the experiments, it was felt that an additional burden on subjects would be excessive and (2) there was no theoretical rationale to expect the act of computing the bonus to affect behavior.

First, we solve for optimal input combinations of x_1 (labor) and x_2 (capital). Parameterizing equation (1), we define firm production as:

$$y^{s,i}=\frac{3(x_1x_2)^{0.4}}{n^{0.2}}, \tag{7}$$

where $n=4$, the number of producers. A well-known characteristic of Cobb-Douglas production is that, for cost minimization, the ratio of the two inputs' marginal products should be equal to the ratios of their relative (after-bonus) prices and also to their relative usage, or:

$$\frac{MPx_1}{MPx_2}=\frac{x_1}{x_2}=\frac{w_1}{(1-\beta)w_2}, \tag{8}$$

which implies $w_1x_1=(1-\beta)w_2x_2$ and total production costs of $2(1-\beta)w_2x_2$. Substituting this last relationship into the production function in equation (7) yields a total cost function of:[8]

$$C(y,x_1,x_2)=2\left(\frac{y}{3}\right)^{1.25}\bullet[(1-\beta)w_1w_2]^{0.5}\bullet n^{0.25}. \tag{9}$$

Totally differentiating equation (9) with respect to output yields each firm's marginal cost function, or

$$MC=\frac{5}{2}y^{0.25}\bullet 3^{-1.25}\bullet[(1-\beta)w_1w_2]^{0.5}\bullet n^{0.25}. \tag{10}$$

Under perfect competition, firms will produce until marginal cost equals the price of output (or $MC=p$). Utilizing this, we solve for aggregate demand (see eqs. [A3] through [A7] in appendix A):

$$Y^D=\left(\frac{16}{625}\right)^{0.2}\bullet 3[(1-\beta)w_1w_2]^{-0.4}\bullet\alpha^{0.8}. \tag{11}$$

We can then substitute equation (11) into a firm's input demand function (which is also defined in terms of w_1, w_2, and y) and, after eliminating common terms and utilizing previous definitions, solve for aggregate market-clearing capital demand (X_2^*) and its price (w_2^*), respectively, as follows (see eqs. [A8] through [A11] in appendix A):

$$w_2^*=\sqrt{\frac{2\alpha}{5b(1-\beta)}}, \tag{12}$$

$$X_2^*=\sqrt{\frac{2\alpha b}{5(1-\beta)}}. \tag{13}$$

Using the same type of analysis, we solve for market-clearing labor quantity and price as:

---

[8] See equations (A1) and (A2) in appendix A.

$$X_1^* = \sqrt{\frac{2\alpha a}{5}},$$

and

$$w_1^* = \sqrt{\frac{2\alpha}{5a}},$$

respectively. We can then substitute these results into the firms' production functions to yield their aggregate production, or

$$Y^* = 3\left(\frac{2\alpha}{5}\right)^{0.4} \bullet (ab)^{0.2} \bullet (1-\beta)^{-0.2}.$$

The market-clearing price for output can be found by substituting w_1^* and w_2^* into the market-clearing demand function (see eq. [A6] in appendix A) to yield the market-clearing output price, or $p^* = 0.48\alpha^{0.6} \bullet (ab)^{0.2}(1-\beta)^{-0.2}$.

Thus, we have solved for all three market-clearing prices and quantities in terms of the exogenous parameters a, b, α, and β. The actual values for the first three parameters (5.12, 0.72, and 500) were chosen because they produced isoquants that resulted in strictly increasing output quantity in decreasing units when using discrete unit inputs from the labor and capital markets. In addition, the parameter values lead to price and quantity predictions in the low double-digit range. Because of the restriction to integer values (the computer software would not allow trading in partial units or prices), slight gaps between supply and demand near the equilibrium values resulted in equilibrium price and quantity ranges.[9] For the benchmark (no tax benefit) scenario, the solutions are $x_1^* \in [28, 32]$; $x_2^* \in [12, 14]$; $y^* \in [32, 33]$; $w_1^* \in [6, 7]$; $w_2^* \in [16, 22]$; and $p^* \in [14, 15]$. In the tax benefit setting, $x_1^* \in [28, 32]$; $x_2^* \in [16, 17]$; $y^* = 36$; $w_1^* \in [6, 7]$; $w_2^* \in [22, 25]$; and $p^* = 13$.

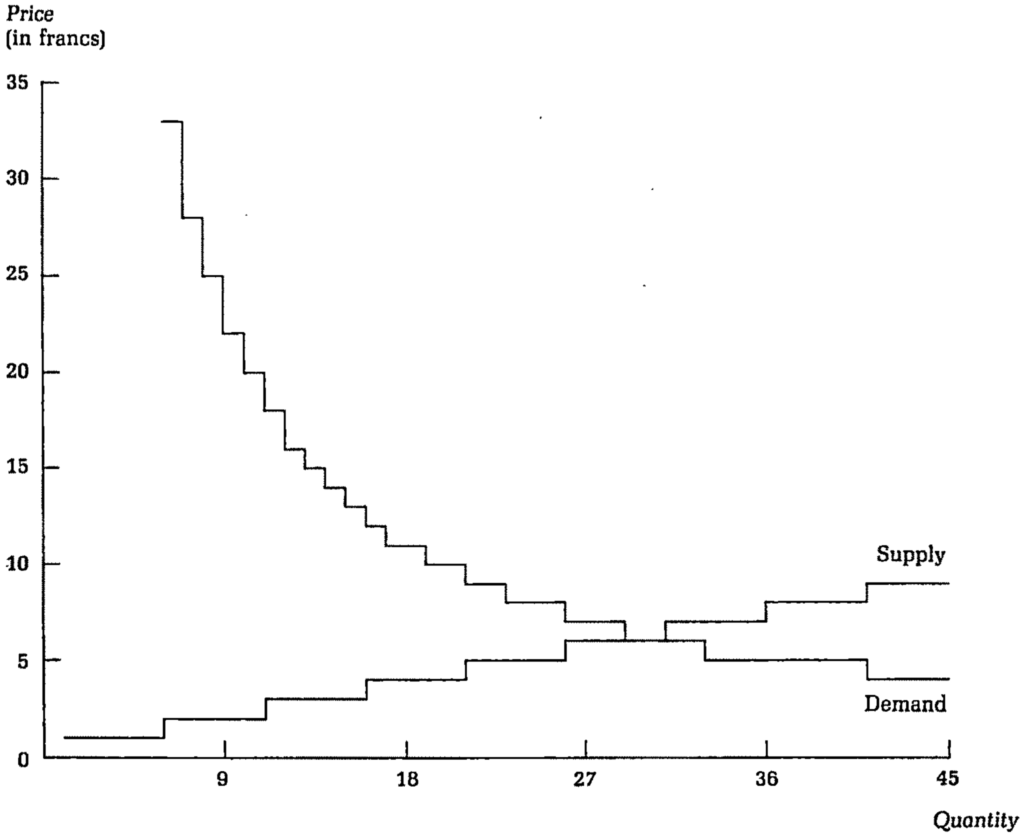
Integer-form demand and supply functions for the labor, capital, and output markets are shown in figures 2 through 4, respectively. Near-equilibrium prices and quantities are shown beneath each figure (i.e., factor suppliers' cost structures and output buyers' redemption values were rounded to integer values such that marginal unit profits were ensured). Integer-form input demand and output supply functions for producers were derived from equations (A9), (A11), and (A12), respectively, in appendix A.

The general result of a tax benefit for capital is an income effect and a substitution effect. The income effect occurs because more purchasing power is available, which results in increased purchases of both labor and capital. The substitution effect results in more capital and less labor purchased. Thus, there are countervailing forces affect-

---

[9] Note that the continuous-form solutions for the benchmark (no tax benefit) setting are: $x_1^* = 32$; $x_2^* = 13.86$; $y^* = 33.42$; $w_1^* = 6.25$; $w_2^* = 19.25$; and $p^* = 15.39$. Similarly, the continuous solutions for the tax benefit setting are: $x_1^* = 32$; $x_2^* = 16.97$; $y^* = 36.24$; $w_1^* = 6.25$; $w_2^* = 23.58$; and $p^* = 13.39$. In addition to supply-demand gaps caused by the requirement of integer solutions, the requirement that (for Cobb-Douglas cost minimization) amounts spent (after bonuses) on labor and capital should be equal resulted in equilibrium quantities of slightly lower levels than the competitive equilibrium solutions (see appendix B for the cost-minimizing solutions). In the above range of solutions, the lower bound represents the cost-minimizing solutions and the upper bound the competitive equilibrium solutions.

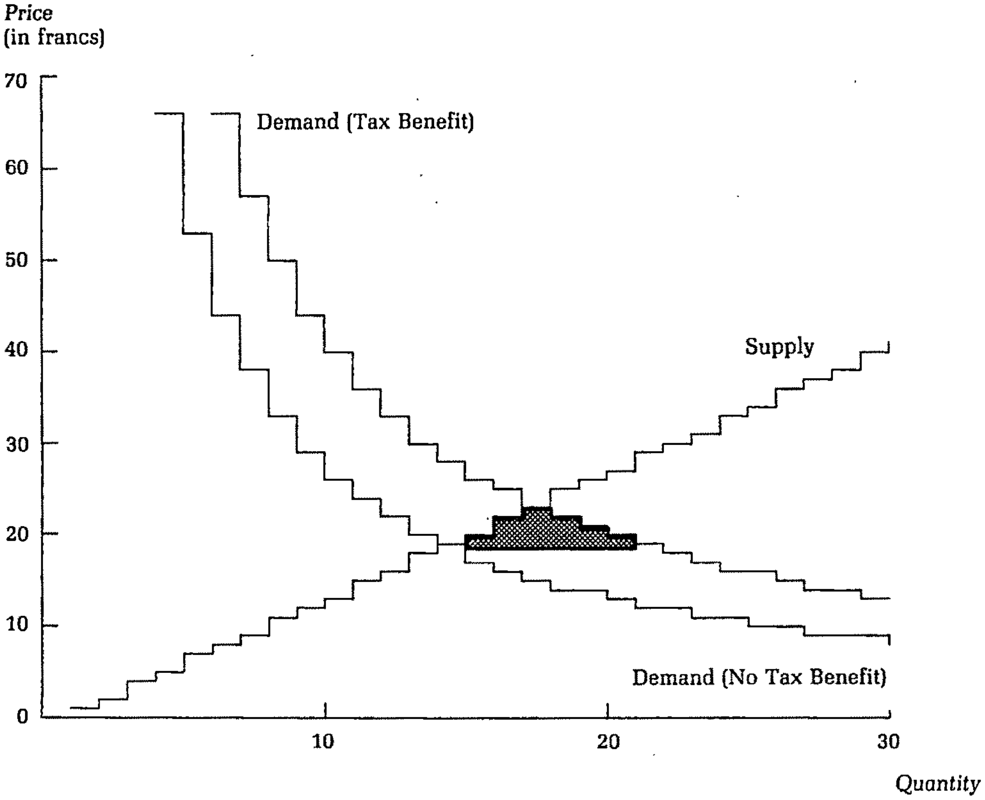
Figure 2
Integer-Form Supply and Demand for the Labor Market



Note: The near-equilibrium prices and quantities (Q) are as follows, with P^S = minimum price demanded by suppliers, and P^D = maximum price payable by producers.

Q	P^S	P^D
24	5	8
25	5	8
26	6	7
27	6	7
28	6	7
29	6	6
30	6	6
31	6	6
32	7	6
33	7	6
34	7	5
35	7	5
36	8	5
37	8	5

Figure 3
Integer-Form Supply and Demand for the Capital Market

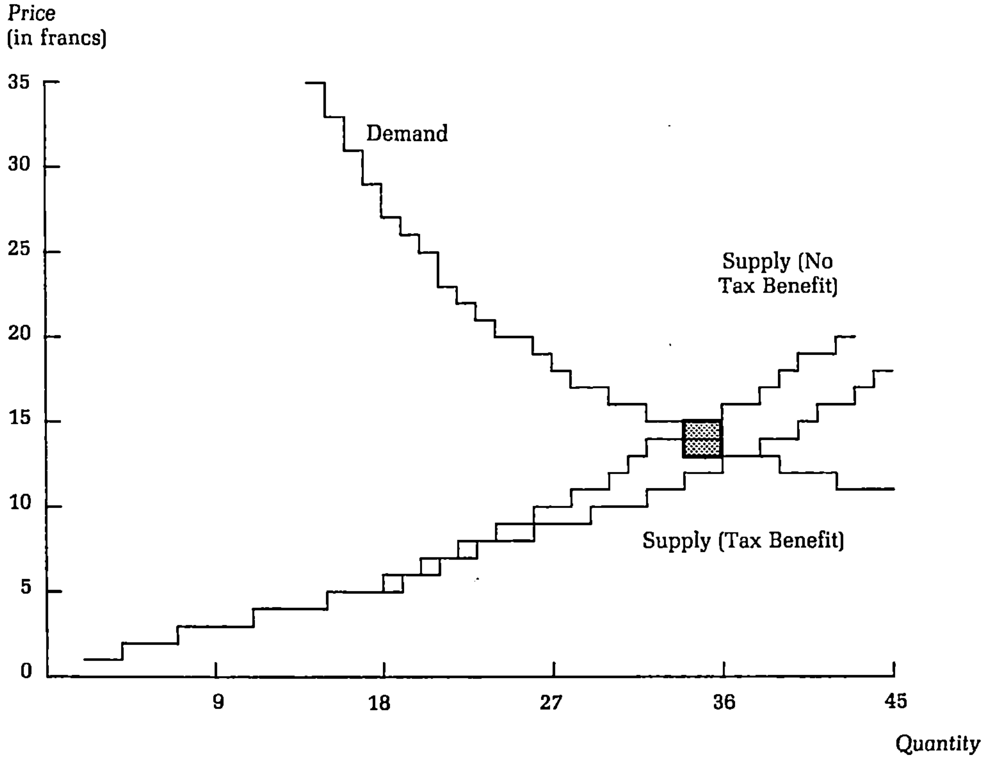


Note: The shaded area in the figure represents "implicit taxes" imposed upon producers because of price increases in the capital market. The near-equilibrium prices and quantities (Q) are as follows, with $P^D_{(T=E)}$ = maximum price payable by producers when no tax benefit, $P^D_{(T>E)}$ = maximum price payable by producers when tax benefit, and P^S = minimum price demanded by suppliers.

Q	$P^D_{(T=E)}$	$P^D_{(T>E)}$	P^S
10	26	40	13
11	24	36	15
12	22	33	16
13	20	30	18
14	19	28	19
15	17	26	20
16	16	25	22
17	15	23	23
18	14	22	25
19	14	21	26

ing labor. In the experiment, the integer solutions of the particular parameters chosen resulted in exactly offsetting income and substitution effects. The interested reader is referred to Henderson and Quandt (1980, chap. 4) for a discussion of income and substitution effects and the firm's production decisions. Thus, figure 2 shows identical labor demand functions across the two tax regimes.

Figure 4
Integer-Form Supply and Demand for the Output Market



Note: The shaded area in the figure represents loss of producers surplus because of price decline in the output market. The near-equilibrium prices and quantities (Q) are as follows, with $P^S_{(T=E)}$ = minimum price demanded by suppliers when no tax benefit, $P^S_{(T>E)}$ = minimum price demanded by suppliers when tax benefit, and P^D = maximum price payable by demanders.

Q	$P^S_{(T=E)}$	$P^S_{(T>E)}$	P^D
26	10	9	14
27	10	9	18
28	11	9	17
29	11	10	17
30	12	10	16
31	13	10	16
32	14	11	15
33	14	11	15
34	15	12	14
35	15	12	14
36	16	13	13
37	16	13	13

Figure 3 shows the CE solutions for capital under the no tax benefit (benchmark) and tax benefit conditions. As depicted, the tax benefit shifts the demand curve for capital, which results in an increase in both quantity and price. Thus, part of the tax subsidy is bid away in the form of higher prices, an effect alternatively referred to as "tax shifting" or implicit taxes.[10] As Scholes and Wolfson (1992, 92–96) point out, if there is any elasticity of supply of depreciable assets, some of the tax shield should be paid away in the form of higher prices paid for assets. Consideration of the implicit taxes attributable to capital assets thus increases the effective tax burden of capital relative to labor. In figure 3, the shaded area can be thought of as implicit taxes from elasticity of supply. Thus, while more capital may be purchased than before as a result of the tax subsidy, less can be purchased than is desired because of the market-price increase.

Figure 4 depicts supply and demand in the output markets. As discussed previously, the tax benefit sessions are predicted to have higher capital purchases that result in an increase in output over the benchmark setting (given no change in labor inputs). This outward shift in the supply function results in greater sales in units, but a lower sales price. The shaded area of the figure represents the resultant loss of the producers' surplus because of such tax shifting, and may be thought of as the component of producers' tax benefits shifted forward to consumers.

The combined effects of the markets on producers are shown hyperspatially in figure 5. The first representation depicts the aggregate integer-form isoquants with areas of equal shading representing roughly equivalent isoquants. The flat, white area at the top represents output that would be sold at a loss, in either experimental condition, given the output market demand schedule. The second and third representations show aggregate integer-form isocosts for the benchmark and tax benefit conditions, respectively. Areas of approximately equal shading represent roughly equal isocosts, with white areas representing net losses to producers, given market prices. We see that the loss frontier occurs at much higher levels of input quantities in the accelerated tax depreciation setting. As can be seen in figure 5, producers face an enormous array of feasible profit-generating input combinations.

Given the foregoing market adjustments, we now return to our original research question, and posit:

H1: With the introduction of a tax benefit (in the form of accelerated depreciation or ITC), the quantity of capital purchased will increase.

H2: With the introduction of a tax benefit (in the form of accelerated depreciation or ITC), the price paid for capital will increase.

Results of the experimental tests of these hypotheses are reported in the next section.

III. Results

Equilibrating Process

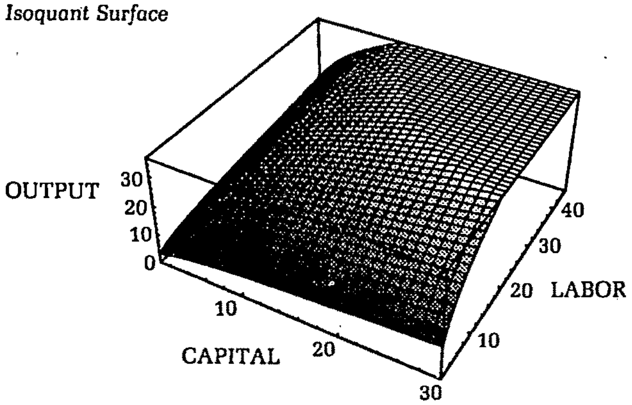
Since the production economy experiment is unique and fairly complex, it is useful to examine prices in the markets to qualitatively see if they converge to equilibrium. Figures 6 through 11 display contract prices for Experiments 1 through 6, respectively.

---

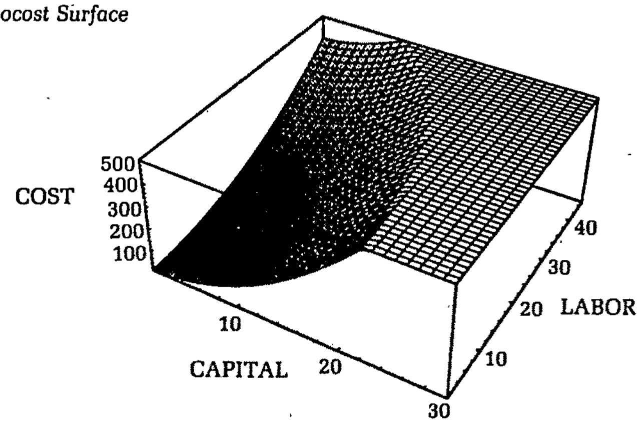
[10] A considerable body of theoretical research has examined shifting of corporate taxes and its long-run equivalent in a general equilibrium setting, tax incidence (most such models have been variants of the original Harberger [1962] study). Empirical studies have been restricted to shifting (which is essentially intermediate-run, i.e., assumed to occur for up to ten years) and have found inconclusive results. Although this research has focused on the shifting of tax costs (as opposed to benefits) and examines shifting of the overall corporate tax burden (as opposed to a tax benefit for only one factor of production), the same principles should apply in the case of tax benefits for depreciable assets. That is, after any tax law change, all markets adjust such that prices and quantities reflect the tax change in the new equilibrim.

Figure 5
Isocost and Isoquant Surfaces

Isoquant Surface



Isocost Surface



Isocost Surface with Tax Subsidy

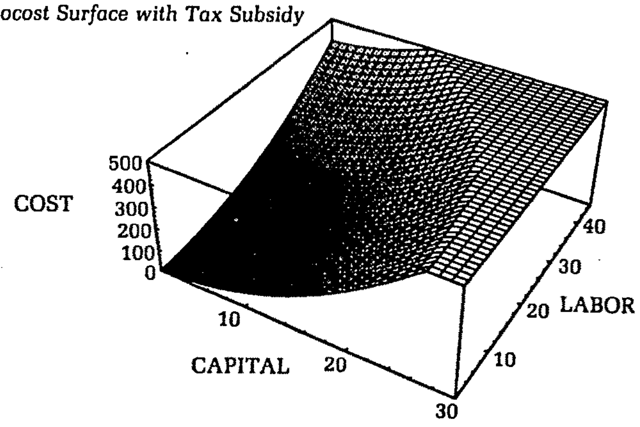
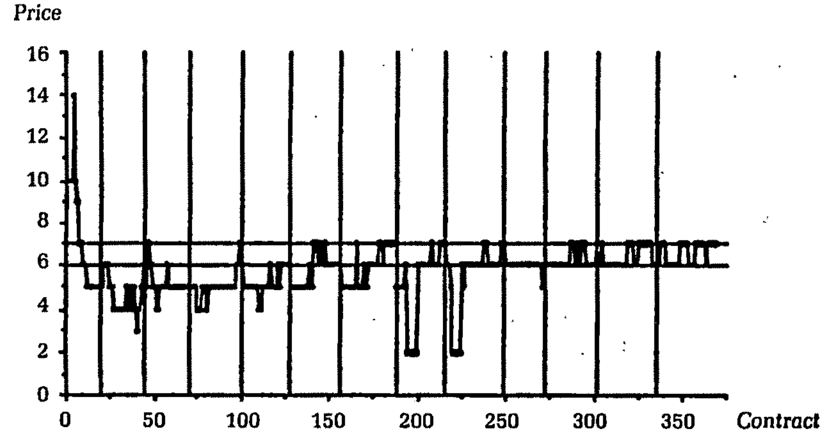
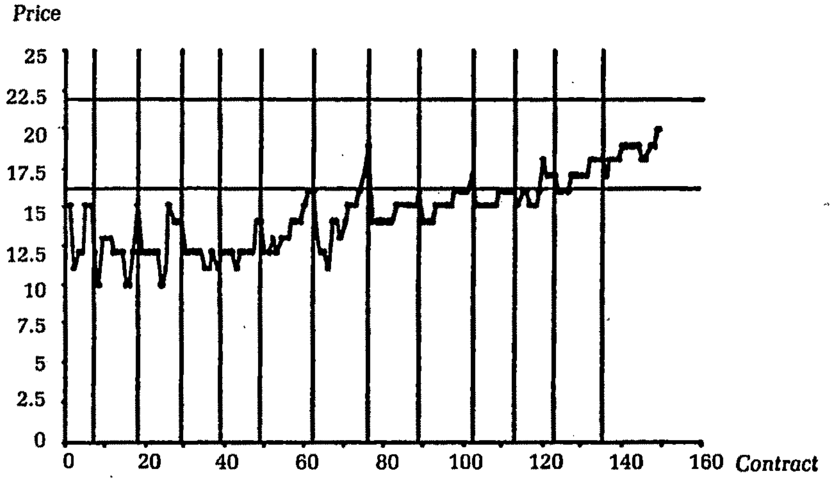


Figure 6
Contract Prices—Experiment 1

Labor Market



Capital Market



Output Market

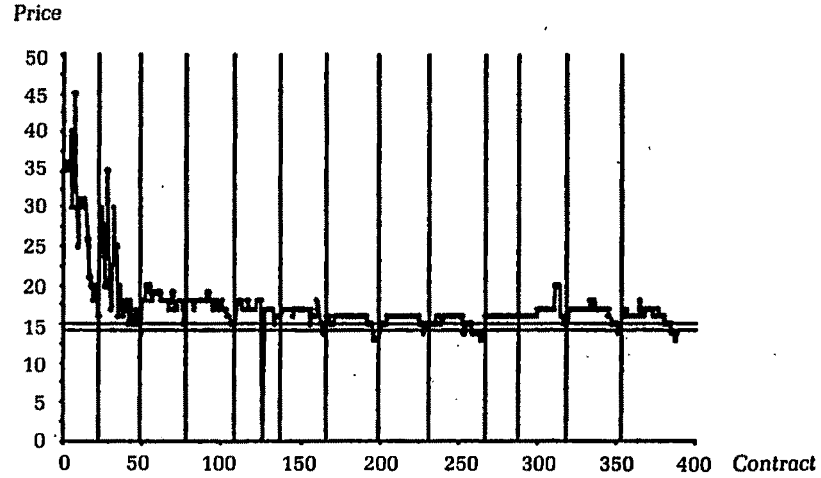


Figure 7
Contract Prices—Experiment 2

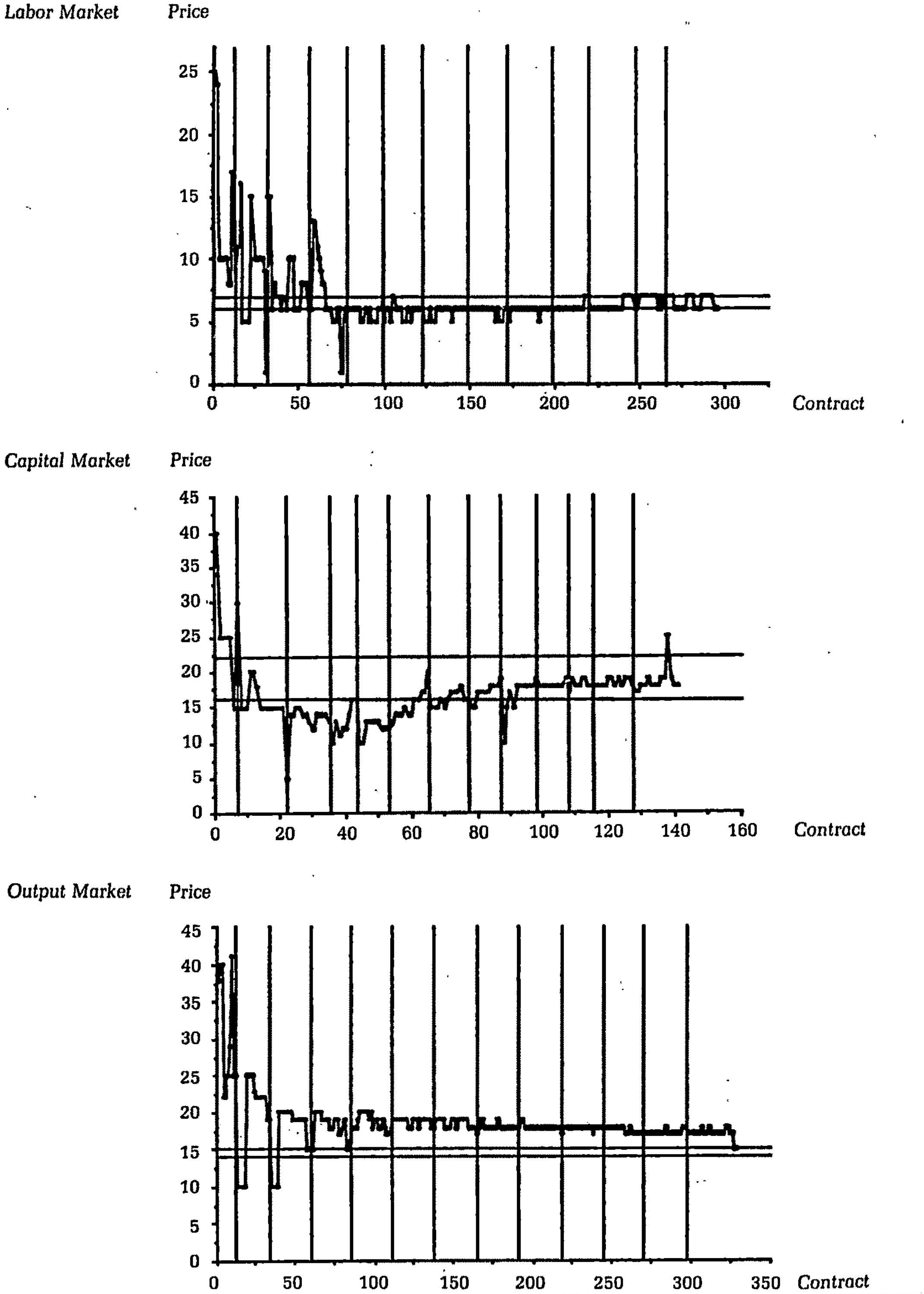
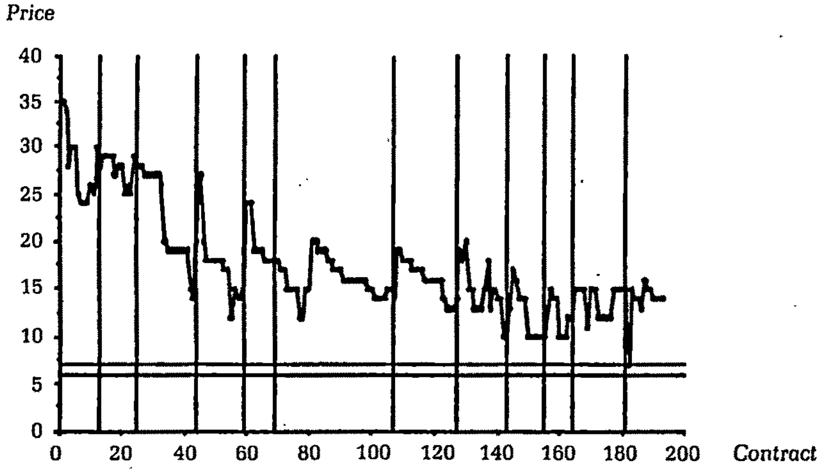
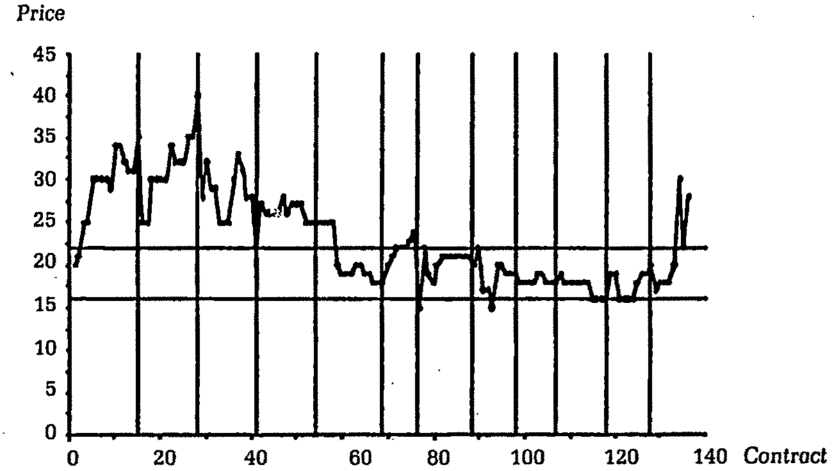


Figure 8
Contract Prices—Experiment 3

Labor Market



Capital Market



Output Market

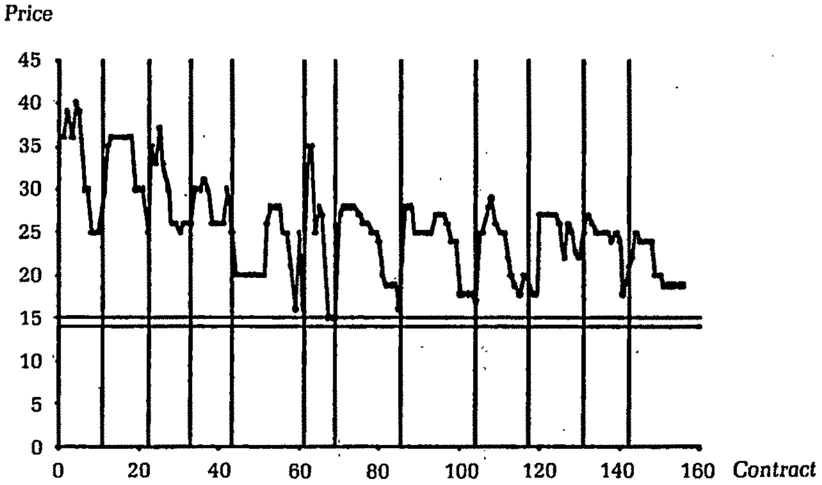


Figure 9
Contract Prices—Experiment 4

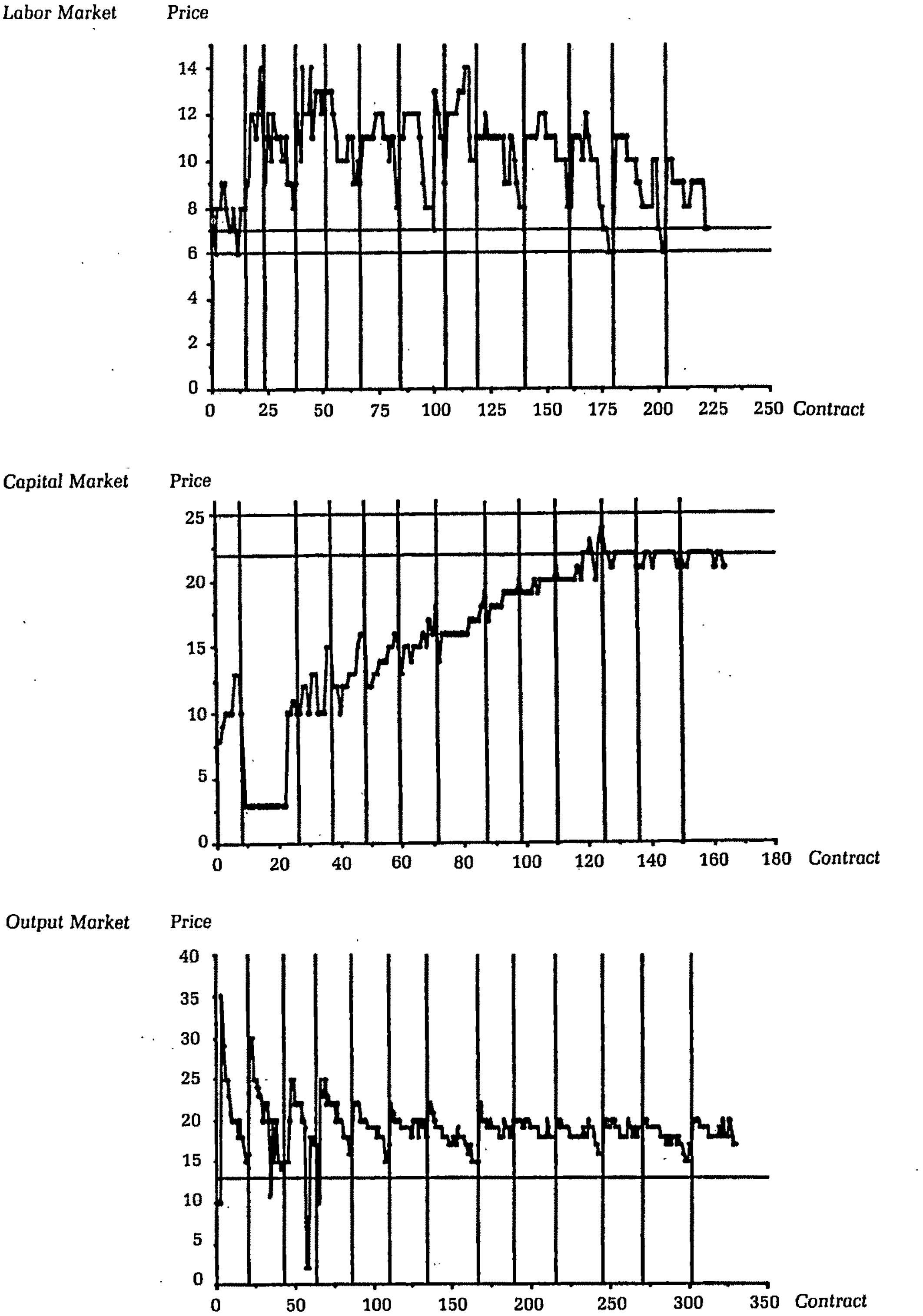


Figure 10
Contract Prices—Experiment 5

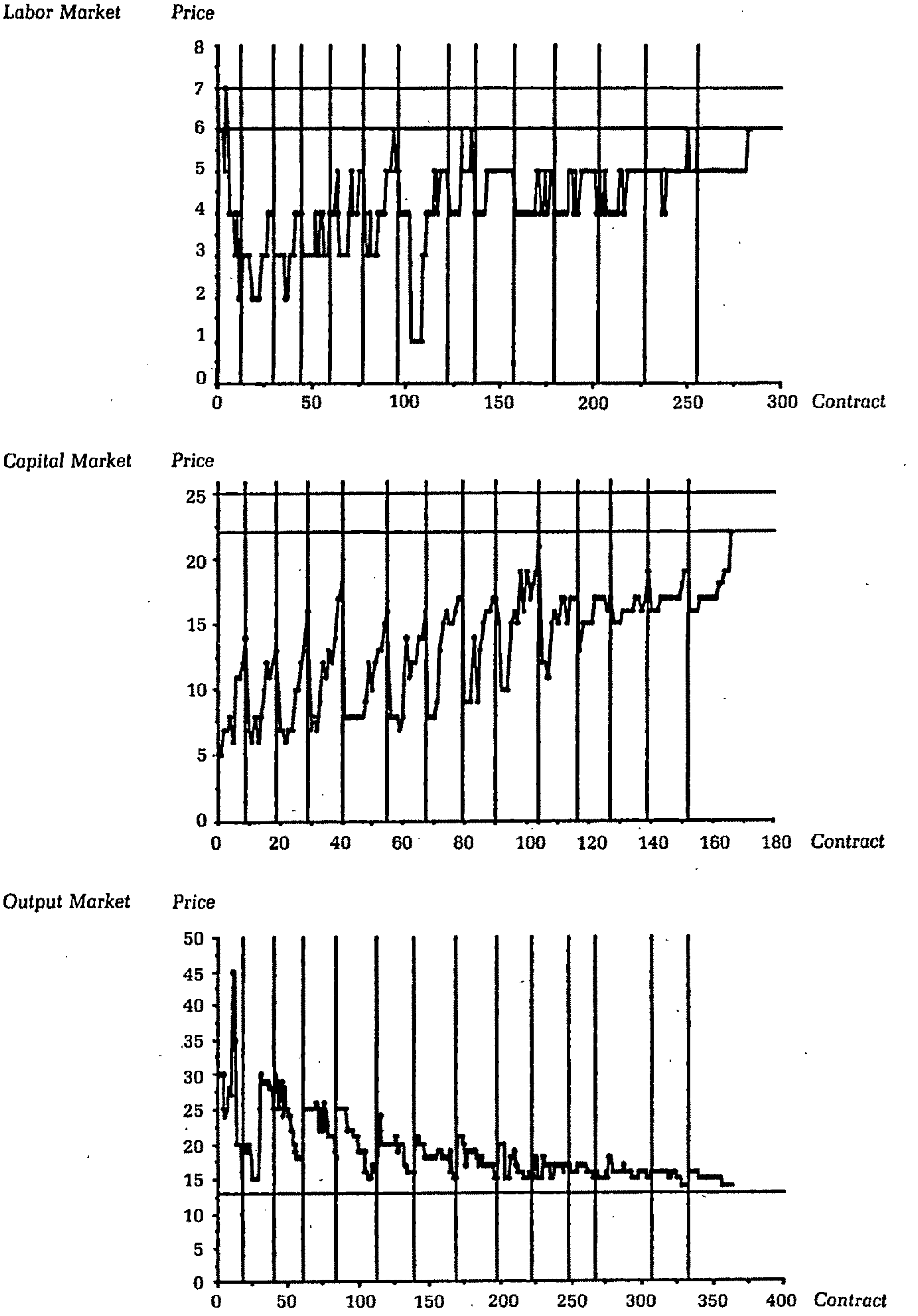
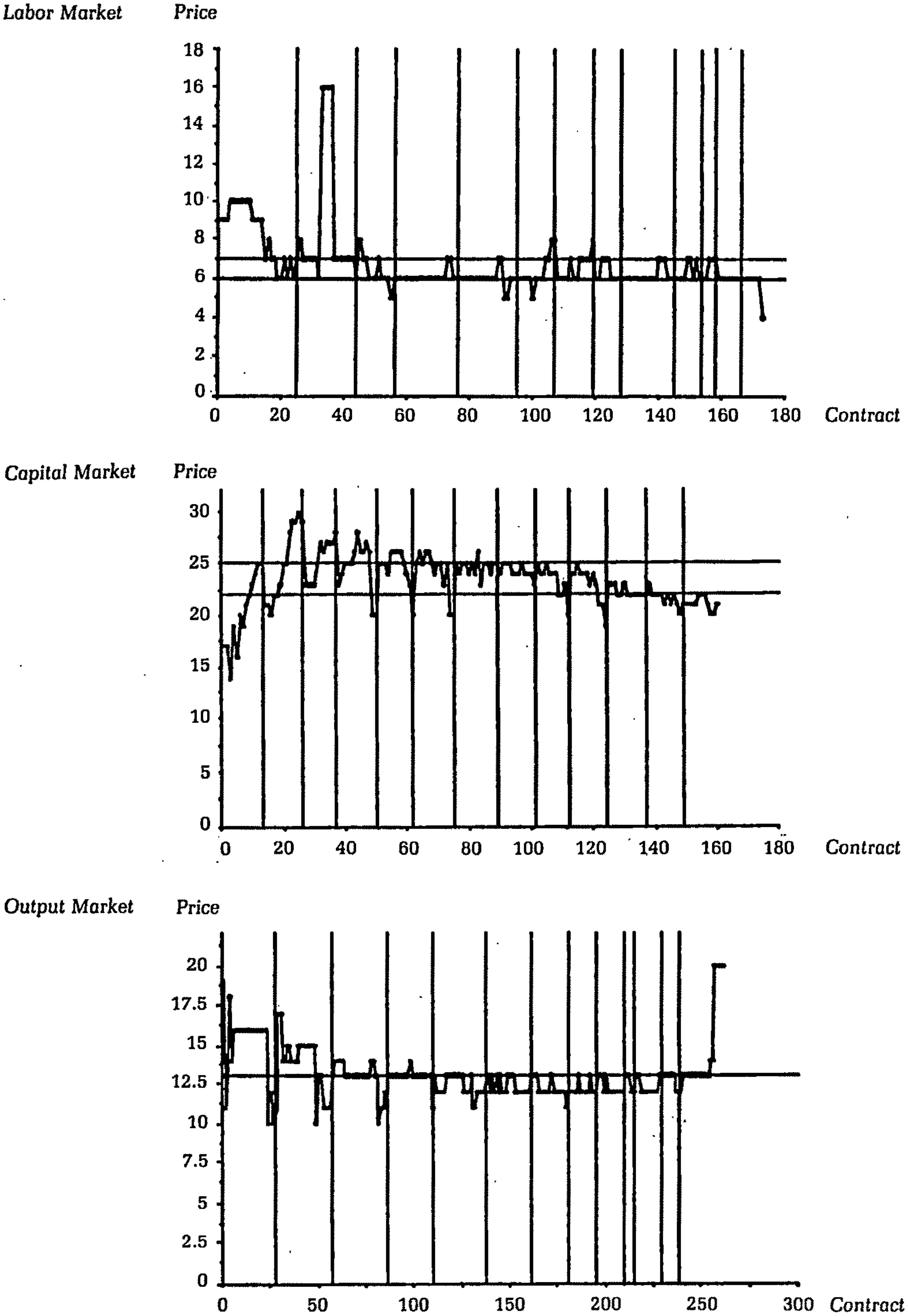


Figure 11
Contract Prices—Experiment 6



Experiments 1 through 3 feature no tax benefit (e.g., tax and economic depreciation equality or no ITC), unlike Experiments 4 through 6, which feature a tax benefit. In each figure, prices for all three markets in a specific experiment are shown. Horizontal lines represent predicted equilibrium price (ranges). Market periods are separated by vertical lines.

As can be seen, prices generally appear to converge toward an equilibrium. For labor markets, prices tend to converge from either below or above the CE. For depreciable asset markets, prices generally converge from below. For ouput markets, prices stabilize quite quickly. This last phenomenon appeared because subjects tended to (on average) sell units at prices sufficient to cover their average costs of production (net of any tax benefits), which they calculated before the output markets opened.

As an alternative to examining contract prices qualitatively, we employed the double exponential adjustment model of Goodfellow and Plott (1990) to estimate equilibrium convergence. The model for any given market is,

$$|P_{tc}-P^*| = \alpha e^{\beta t+\gamma c}, \tag{14}$$

where t is the period number, c is the contract (or transaction) number in the period, and P^* is the competitive equilibrium price. Contract prices are converging to the CE across periods if $\beta<0$ and within periods if $\gamma<0$ (the intercept, α, simply denotes the distance from equilibrium at the beginning of the experiment). Regression results for all markets and both tax regimes are shown in tables 1 and 2. In table 1, equilibrium convergence to the upper bound of CE prices is measured; in table 2, the convergence to the lower bound is measured. As can be seen, virtually all of the markets are equilibrating both within and across periods; 55 out of 66 (or 83 percent) of the β and γ coefficients are significant (at 5 percent) and negative. Of the other 11, only three are positive and significant. As can be seen by mean price data in tables 3 and 4 (discussed below), this is because prices very quickly achieved near-CE levels, so that no convergence was observed.

Quantities and mean prices for each of the benchmark (no tax benefit) experiments are shown in table 3 by market and period. Equivalent data for the tax benefit experiments is shown in table 4. Much like the price data, the quantities exhibit differences across experiments; four experiments show a pattern of increasing transactions as the experiment progresses, but two exhibit relatively stable quantity behavior. What is striking is that most experiments seem to immediately (and continuously) have approximately the correct proportion of capital-to-labor mixes (about 1:2), although there is often variance from the correct absolute quantities. Both tables show system efficiency, operationally measured as the sum of all agents' actual profits divided by the maximum possible profits at CE prices. For all but Experiment 3, relatively high efficiency occurred, with typical measures above 95 percent for the later half of each experiment. Note that for many periods, a small quantity of inventory was unsold. This generally resulted (given downward-sloping output demand curves) in observed output prices that appeared high for the output produced.

Hypothesis Tests

An inspection of the data in tables 3 and 4 leaves the impression that, while capital market prices are generally higher in the tax benefit markets, quantities of capital purchased are generally equivalent with the benchmark condition. A comparison of aver-

Table 1
Regression Results for the Highest Equilibrium Point

Model: $|P_{tc} - P^*| = \alpha e^{\beta t + \gamma c}$.

Experiment	*Market*	α	β	γ	n	R^2	*Durbin-Watson Test*
No Tax Benefit: (benchmark)							
1	Labor	3.42	−0.08	−0.03	369	0.30	0.62
		(15.31)	(−8.39)	(−6.49)			
	Capital	13.22	−0.06	−0.02	149	0.70	0.61
		(31.44)	(−15.62)	(−5.95)			
	Output	46.27	−0.58	−0.07	388	0.76	1.26
		(14.52)	(−15.59)	(−15.71)			
2	Labor	14.94	−0.32	−0.08	295	0.47	0.85
		(7.21)	(−8.73)	(−5.47)			
	Capital	9.97	−0.06	−0.02[a]	141	0.31	1.00
		(12.97)	(−7.38)	(−1.55)			
	Output	15.36	−0.17	−0.03	332	0.56	0.59
		(19.31)	(−14.31)	(−8.34)			
3	Labor	29.94	−0.13	−0.03	193	0.80	0.73
		(29.79)	(−23.24)	(−10.67)			
	Capital	6.33	−0.12	0.04[b]	136	0.31	0.82
		(7.69)	(−5.73)	(3.65)			
	Output	25.37	−0.08	−0.06	156	0.54	0.74
		(16.57)	(−9.38)	(−7.95)			
Tax Benefit:							
4	Labor	4.94	−0.01[a]	−0.04	223	0.23	0.53
		(14.97)	(−0.62)	(−7.42)			
	Capital	24.05	−0.16	−0.01[a]	164	0.84	0.45
		(25.82)	(−24.95)	(−1.49)			
	Output	10.95	−0.03	−0.03	330	0.38	0.91
		(23.58)	(−6.09)	(−12.10)			
5	Labor	4.35	−0.04	−0.01	283	0.32	0.50
		(26.21)	(−8.80)	(−4.49)			
	Capital	24.28	−0.06	−0.05	166	0.82	0.98
		(39.66)	(−20.01)	(−16.21)			
	Output	19.28	−0.14	−0.02	363	0.60	0.45
		(21.99)	(−18.35)	(−6.82)			
6	Labor	2.63	−0.13	−0.03	173	0.08	0.63
		(4.59)	(−3.20)	(−2.15)**			
	Capital	1.60	0.01[a]	0.04[b]	160	0.02	0.562
		(3.21)	(0.33)	(1.67)*			
	Output	4.86	−0.45	−0.01[a]	261	0.20	0.473
		(4.20)	(−3.99)	(−1.076)			

Note: The model is defined in the text. The number of observations is denoted by n.

[a] Not significantly different from zero at the 5 percent level with a two-tailed test.
[b] Significantly positive at the 5 percent level.

* $p < 0.10$.
** $p < 0.05$. All other values are significant at $p < 0.01$.

Table 2
Regression Results for the Lowest Equilibrium Point

Model: $|P_{tc} - P^*| = \alpha e^{\beta t + \gamma c}$.

Experiment	*Market*	α	β	γ	*n*	R^2	*Durbin-Watson Test*
No Tax Benefit: (benchmark)							
1	Labor	2.63 (8.67)	−0.12 (−7.00)	−0.03 (−4.26)	369	0.18	0.61
	Capital	5.74 (10.82)	−0.10 (−8.04)	−0.04 (−3.19)	149	0.37	0.69
	Output	34.16 (13.27)	−0.38 (−12.72)	−0.06 (−13.29)	388	0.70	0.91
2	Labor	26.87 (8.05)	−0.55 (−10.67)	−0.10 (−6.17)	295	0.58	0.97
	Capital	9.62 (4.09)	−0.15 (−3.62)	−0.08 (−2.70)	141	0.17	1.27
	Output	13.45 (18.78)	−0.12 (−12.73)	−0.02 (−6.47)	332	0.45	0.53
3	Labor	30.24 (31.34)	−0.12 (−23.20)	−0.03 (−10.62)	193	0.80	0.72
	Capital	15.77 (12.06)	−0.20 (−12.75)	0.02[b] (3.00)	136	0.67	0.64
	Output	25.85 (17.92)	−0.07 (−9.49)	−0.05 (−8.09)	156	0.54	0.73
Tax Benefit:							
4	Labor	5.86 (17.87)	−0.00[a] (−0.33)	−0.04 (−7.62)	223	0.24	0.49
	Capital	22.03 (19.49)	−0.22 (−21.24)	0.00[a] (−0.79)	164	0.82	0.42
5	Labor	3.50 (18.74)	−0.06 (−8.97)	−0.02 (−4.36)	283	0.34	0.51
	Capital	22.35 (30.83)	−0.07 (−18.64)	−0.07 (−14.88)	166	0.81	0.93
6	Labor	7.77 (4.70)	−0.53 (−5.38)	−0.05 (−3.73)	173	0.30	0.66
	Capital	3.81 (8.60)	−0.10 (−7.73)	0.02[a] (1.63)	160	0.31	0.81

Note: The model is defined in the text. The number of observations is denoted by *n*.

[a] Not significantly different from zero at the 5 percent level with a two-tailed test.
[b] Significantly positive at the 5 percent level.

* $p < 0.10$.
** $p < 0.05$. All other values are significant at $p < 0.01$.

Table 3
Mean Prices, Volume, and System Efficiency for the Benchmark Markets

Experiment	*Period*	P^L	Q^L	P^K	Q^K	P^O	$Q^{O.P}$	$Q^{O.S}$	*System Efficiency*
1	1	7.16	18	13.14	6	28.46	22	21	0.90
	2	4.62	25	12.00	10	19.96	29	26	0.67
	3	5.16	24	12.55	10	18.14	28	27	0.92
	4	5.03	29	11.70	9	17.40	30	29	0.89
	5	5.22	26	12.30	9	16.48	29	28	0.94
	6	5.79	27	13.62	12	16.45	30	28	0.95
	7	5.94	31	14.29	13	15.47	35	33	0.99
	8	5.11	27	14.62	12	15.66	32	31	1.00
	9	5.38	33	15.23	12	15.11	36	34	1.00
	10	5.95	21	15.46	10	16.00	22	20	0.96
	11	6.27	29	16.20	9	16.70	30	30	1.00
	12	6.35	33	17.08	11	16.59	33	33	0.99
	13	6.56	33	18.57	13	15.89	36	35	1.00
2	1	14.00	12	26.43	6	30.39	14	12	0.80
	2	9.06	17	15.20	14	19.48	24	20	0.91
	3	8.00	24	13.85	12	16.93	27	26	0.99
	4	6.76	20	13.00	7	18.42	25	23	0.99
	5	5.73	21	12.10	9	18.77	28	25	0.96
	6	5.83	23	15.33	11	18.78	30	26	0.98
	7	5.73	25	16.08	11	18.44	29	26	0.97
	8	5.83	22	17.10	9	18.19	27	25	0.95
	9	5.92	25	16.73	10	18.04	28	27	0.97
	10	6.14	21	18.00	9	17.96	26	25	0.98
	11	6.27	25	18.43	6	17.56	26	24	0.98
	12	6.80	19	18.42	11	17.19	27	26	0.98
	13	6.52	28	18.64	13	17.06	32	31	1.00
3	1	27.92	39	29.13	21	32.27	43	16	0.61
	2	27.67	30	31.54	22	33.27	37	18	0.64
	3	21.44	42	28.08	17	29.36	41	20	0.74
	4	17.27	27	26.23	16	28.00	33	21	0.77
	5	17.47	28	20.60	15	22.11	32	19	0.86
	6	16.28	29	22.00	11	24.38	30	14	0.58
	7	16.15	24	20.08	16	24.06	34	24	0.76
	8	14.69	28	18.67	14	23.84	33	26	0.80
	9	12.33	31	18.22	19	23.23	38	24	0.73
	10	12.22	17	17.36	12	24.29	25	23	0.83
	11	13.88	32	17.80	21	24.09	40	20	0.69
	12	13.75	35	21.38	18	21.21	38	26	0.76
Competitive Equilibrium (CE)		6-7	28-29	16-22	12-14	14-15	32-33	32-33	100

Note: The benchmark markets have no tax benefit. P^L and Q^L are price and quantity of labor; P^K and Q^K are price and quantity of capital; P^O is price of output; $Q^{O.P}$ is quantity of output produced; and$Q^{O.S}$ is quantity of output sold.

ages across the experimental conditions confirms this suspicion. If we use the averages of the last three periods across experiments, average capital purchased per period are 12.33 units at 20.47 francs in conditions with tax benefits and 12.67 units at 18.21 francs in the benchmark conditions. Similarly, averages across the last five periods yield per-period average capital purchases of 12.4 units at 19.94 francs in the tax benefit experi-

Table 4
Mean Prices, Volume, and System Efficiency for the Tax Benefit Markets

Experiment	*Period*	P^L	Q^L	P^K	Q^K	P^O	$Q^{O,P}$	$Q^{O,S}$	*System Efficiency*
4	1	7.67	14	10.38	7	20.40	20	19	0.89
	2	11.11	8	4.61	17	20.39	25	22	0.93
	3	10.29	13	11.82	10	17.67	23	20	0.93
	4	12.39	12	13.09	10	20.23	21	21	0.92
	5	10.69	15	13.91	10	19.00	24	23	0.95
	6	10.82	16	15.33	11	19.54	24	23	0.94
	7	10.52	20	16.44	15	17.91	32	31	0.99
	8	12.07	13	18.55	10	19.35	23	22	0.95
	9	10.29	20	19.67	11	18.96	27	25	0.98
	10	10.62	20	21.13	14	18.33	30	29	1.00
	11	9.21	18	21.82	10	19.00	27	24	0.96
	12	9.13	22	21.71	13	17.97	31	30	1.00
	13	8.79	18	21.71	13	18.69	29	28	1.00
5	1	4.42	11	9.00	8	27.28	19	17	0.75
	2	2.88	16	9.30	9	22.64	23	21	0.93
	3	3.13	14	9.50	9	23.40	21	19	0.92
	4	3.33	14	11.64	10	23.46	23	23	0.89
	5	4.00	17	10.40	14	20.36	30	27	0.94
	6	4.21	18	11.00	11	19.11	29	26	0.94
	7	3.50	25	13.08	11	18.27	30	29	0.95
	8	4.71	13	13.09	10	17.93	29	28	0.98
	9	4.71	20	15.71	13	15.88	28	24	0.94
	10	4.23	21	15.08	11	16.64	29	24	0.99
	11	4.52	22	15.82	10	15.93	29	27	0.95
	12	4.56	24	16.25	11	15.87	32	29	0.98
	13	5.00	27	17.15	12	15.39	33	25	1.00
	14	5.07	27	17.57	13	15.00	34	30	1.00
6	1	8.28	29	20.15	13	15.11	34	31	0.66
	2	8.84	33	24.92	13	13.83	34	34	0.91
	3	6.33	28	25.36	11	12.97	30	29	0.90
	4	6.10	24	24.69	13	12.96	30	28	0.92
	5	6.00	27	24.58	13	12.37	32	32	0.95
	6	6.42	25	24.23	13	12.48	30	30	0.93
	7	6.58	34	24.64	14	12.15	36	35	0.95
	8	6.33	25	24.25	14	12.14	32	32	0.95
	9	6.18	25	23.36	14	12.29	32	31	0.95
	10	6.38	27	23.00	15	12.50	33	33	0.95
	11	6.60	21	22.31	13	12.29	29	29	0.92
	12	6.00	26	21.58	13	12.70	32	32	0.94
	13	5.71	27	21.09	13	14.68	32	22	0.87
Competitive Equilibrium (CE)		6-7	28-29	22-25	16-17	13	36	36	100

Note: P^L and Q^L are price and quantity of labor; P^K and Q^K are price and quantity of capital; P^O is price of output; $Q^{O,P}$ is quantity of output produced; and $Q^{O,S}$ is quantity of output sold.

ments and 12.53 units at 17.75 francs for the benchmark experiments. To test for differences between means, we employ a standard two-sample *t*-test.[11] The difference in mean quantity of capital purchased is not significantly different between conditions for both the three- and five year horizon ($t=0.216$ and 0.119, respectively). However, the difference in mean contract price is significantly different ($t=2.06$, $p=0.03$ for the three-year horizon; $t=2.681$, $p=0.007$ for the five-year horizon).

---

[11] Note that, since observations lack independence within groups for the *t*-test, some caution regarding reliance on these results is in order.

Thus, there is support for our hypothesis of capital price increases (or implicit taxes) in response to the introduction of a tax benefit. Capital prices are, on average, 12 percent higher in the tax benefit experiments. However, there is no support for the hypothesized increase in the quantity of capital purchased.[12] This latter result is partly driven by Experiment 3 in the benchmark condition, where capital purchased is much too high. After excluding this experiment, capital purchased in the tax benefit manipulation is uniformly higher (by about 8 percent in the last three periods) than the benchmark capital purchases. However, even this increase is too low; our equilibrium predictions are that aggregate capital purchased should increase by at least 14 percent (or by 2 to 4 units) even after taking into account the demand-dampening effect of implicit taxes. This result begs a closer inspection of the data in terms of specific point predictions, which is addressed in the next section.

Point Predictions

Results by period from tables 3 and 4 are combined into averages in table 5, which reports mean contract prices and quantities by experimental condition, experiment, and market. Although most of the experiments achieved high efficiency halfway through the experiment, exact (and stable) CEs were not achieved. Thus, there is no obvious point at which to begin examination of the data in terms of point predictions. Accordingly, table 5 reports averages over the last five periods, the last three periods, and for the final period. Since CEs are not achieved, another issue is the range within which actual prices and quantities can be considered "close enough" to predictions. In the only published study of multiple markets with production, Goodfellow and Plott (1990, 975) conclude that final-period prices and quantities within 12 percent of CE "compare favorably" to predictions.[13] Given the greater complexity of the present experiment, we assume that a range of 15 percent is reasonable here, although the use of any range is admittedly *ad hoc*.

For Experiment 1, all of the prices and quantities for the three time-period comparisons are within the acceptable range. For Experiment 2, 94 percent are within the range. Experiment 3, the last of the benchmark experiments, was less successful, with only 39 percent of prices and quantities within the acceptable range. The three experiments combined were, on average, within acceptable price and quantity prediction ranges 78 percent of the time, which is comparable to the results of Goodfellow and Plott (1990) (12 out of 16 prices and quantities for the last period, for four experiments each having two markets). Despite the weaker results in our third experiment, the aggregate results for this regime appear highly encouraging. Recall that producers faced no explicit demand for input or supply of output induced-value schedules, but

---

[12] In contrast with this finding, predicted output prices and all predicted quantities were achieved exactly for the last several periods on almost all contracts in a pilot study with fixed prices for factor inputs and outputs from production, in conformance with the neoclassical "price-taker" assumption.

[13] Prices and quantities in the study by Goodfellow and Plott (1990) were off by 11 percent and 12.5 percent, on average, respectively, and prices in some experiments were off by more than 18 percent. See the conclusion for a further discussion.

Table 5
Mean Quantities and Prices

Experiment and Periods	Labor		Capital		Output	
	Price	Quantity	Price	Quantity	Price	Quantity
No Tax Benefit (benchmark)						
Experiment 1						
Last 5 Periods	6.10[a]	29.80[a]	16.51[a]	11.00[a]	16.06[a]	31.40[a]
Last 3 Periods	6.39[a]	31.67[a]	17.28[a]	11.00[a]	16.39[a]	33.00[a]
Last Period	6.56[a]	33.00[a]	18.57[a]	13.00[a]	15.89[a]	35.00[a]
Experiment 2						
Last 5 Periods	6.33[a]	23.60[a]	18.04[a]	9.80[a]	17.56	27.80[a]
Last 3 Periods	6.53[a]	24.00[a]	18.50[a]	10.00[a]	17.27[a]	28.30[a]
Last Period	6.52[a]	28.00[a]	18.64[a]	13.00[a]	17.06[a]	31.00[a]
Experiment 3						
Last 5 Periods	13.37	28.60[a]	18.69[a]	16.80	23.33	34.80[a]
Last 3 Periods	13.38	28.00[a]	18.85[a]	17.00	23.20	34.30[a]
Last Period	13.75	35.00	21.38[a]	18.00	21.21	26.00
Competitive Equilibrium (CE)	6–7	28–29	16–22	12–14	14–15	32
Tax Benefit						
Experiment 4						
Last 5 Periods	9.61	19.60	21.20[a,b]	12.20[b]	18.59[b]	28.80[b]
Last 3 Periods	9.04[b]	19.30	21.75[a,b]	12.00[b]	18.56[b]	29.00[b]
Last Period	8.79[b]	18.00	21.71[a,b]	13.00[b]	18.69[b]	28.00[b]
Experiment 5						
Last 5 Periods	4.68[b]	24.20[a,b]	16.37[b]	11.40[b]	15.77[b]	31.40[a,b]
Last 3 Periods	4.88[b]	26.00[a,b]	16.99[b]	12.00[b]	15.42[b]	33.00[a]
Last Period	5.07[a,b]	27.00	17.57[b]	13.00[b]	15.00[b]	34.00[a]
Experiment 6						
Last 5 Periods	6.17[a,b]	25.20[a,b]	22.27[a,b]	13.60[a,b]	12.89[a,b]	31.60[a,b]
Last 3 Periods	6.10[a,b]	18.00	22.66[a,b]	13.00	13.22[a,b]	31.00[a,b]
Last Period	5.71[a,b]	27.00[a,b]	21.09[a,b]	13.00[b]	14.68[a,b]	32.00[a,b]
Competitive Equilibrium (CE)	6–7	28–29	22–25	16–17	13	36
Competive Equilibrium (at Lower Isoquant)	5–8	24	16–33	12	9–17	28

[a] Observation is within 15 percent of competitive equilibrium.
[b] Observation is within 15 percent of the first cost-minimizing isoquant below competitive equilbrium.

instead had to derive these intuitively (in conformity with Cobb-Douglas cost minimization and by impounding an asset structure into the value of capital purchases) while interacting in three markets. Since there are several hundred possible interior solution combinations of market prices and quantities (see figure 5), the results here thus are unlikely to be attributable to chance.

For the tax benefit experiments, the results in table 5 appear less encouraging. An inspection of the combined results shows that only 24 of 48 (or 50 percent) of prices and quantities are within acceptable range of point predictions. One striking observation is that, for all markets and experiments, actual quantities are lower than predictions (i.e., pro-

Table 6
Additional Analysis of Tax Benefit Sessions: Actual (Predicted) Mean Prices and Quantities

	Labor		*Capital*		*Output*	
Experiment	*Quantity*	*Price*	*Quantity*	*Price*	*Quantity*	*Price*
Experiment 4						
Period 9	20 (20)*	10.29 (4–9)	11 (11)*	19.67 (15–36)*	27	18.96 (9–18)
Period 10	20 (24)	10.62 (4–9)	14 (13)*	21.13 (19–28)*	30	18.33 (10–16)
Period 11	18 (20)	9.21 (4–10)*	10 (11)*	21.82 (13–40)*	27	19.00 (9–18)
Period 12	22 (28)	9.13 (4–9)*	12 (15)	21.71 (18–30)*	31	17.97 (10–16)
Period 13	18 (24)	8.79 (4–10)*	13 (12)*	21.71 (18–30)*	29	18.69 (10–17)
Experiment 5						
Period 10	21 (24)	4.23 (4–9)*	11 (12)*	15.08 (15–36)*	29	16.64 (10–17)*
Period 11	22 (24)	4.52 (4–9)*	10 (12)	15.82 (13–40)*	29	15.93 (10–17)*
Period 12	24 (26)	4.56 (5–8)*	11 (14)	16.25 (15–36)*	32	15.87 (11–15)
Period 13	27 (28)*	5.00 (6–7)	12 (14)	17.15 (16–33)*	33	15.39 (11–15)*
Period 14	27 (28)*	5.07 (6–7)	13 (15)	17.57 (18–30)*	34	15.00 (12–14)
Experience 6						
Period 9	25 (26)*	6.18 (5–8)*	14 (14)*	23.36 (19–28)*	32	12.29 (11–15)*
Period 10	27 (28)*	6.38 (6–7)*	15 (14)*	23.00 (20–26)*	33	12.50 (11–15)*
Period 11	21 (24)	6.60 (4–9)*	13 (12)*	22.31 (18–30)*	29	12.29 (10–17)*
Period 12	26 (26)*	6.08 (6–7)*	13 (14)*	21.58 (18–30)*	32	12.70 (11–15)*
Period 13	27 (26)*	5.71 (6–7)*	13 (14)*	21.09 (18–30)*	32	14.68 (11–15)*

An * indicates actual price (rounded to nearest integer) within predicted price range, or actual quantity within one unit of predicted quantity. Predicted price ranges, given actual quantities, appear next to actual quantities in parentheses. Predicted and optimal labor and capital inputs, given actual outputs, appear next to actual inputs in parentheses.

ducers settled on production at a lower-than-predicted isoquant). Appendix B shows that, for this tax regime, the next lower Cobb-Douglas cost-minimizing isoquant is the aggregate purchase of 24 units of labor and 12 units of capital, which should yield (at predicted market prices) total after-tax-subsidy factor costs of 156 and 147, respectively. From appendix B, we see that, for this isoquant, predicted quantities (price ranges) for the three markets are 24 (5 through 8), 12 (16 through 33) and 28 (9 through 17).[14] Comparing actual mean prices and quantities to these alternative predictions, with observations falling within 15 percent of predictions indicated in the table by a cross, one can see in Experiments 4 through 6 that 78 percent, 83 percent, and 94 percent of observed prices and quantities are near expected levels. These levels represent an 85 percent accuracy of prediction rate on a combined basis.

These results provide a general picture of what occurred in the tax benefit sessions, but a more detailed analysis may provide additional insight. A period-by-period analysis of each session by market for the final five periods is shown in table 6. Actual prices are shown, as are predicted prices (price ranges), given the actual quantities. Also shown are predicted labor and capital quantities purchased, given the actual outputs produce. Based on lowest-cost labor-capital mixes at the actual output isoquant and allowing for nonsymmetry across producers, these predictions may be thought of as a partial equilibrium analysis of the input market when compared to actual labor and capital.

---

[14] Note that the price range is due to a gap between supply and demand prices at these quantities, i.e., these are not quite market-clearing prices and quantities.

Asterisks in table 6 indicate that observed prices fall within the predicted price ranges, or that observed labor or capital are within one unit of predictions. For the price predictions, 33 of the 45 predictions (or 75 percent) are supported. Recall that producers had no explicit input demand schedules or output price schedules, but entered into contracts on the basis of observed prices and their own induced input demand and output supply functions. The price results thus appear highly encouraging. It is also worth noting that observed capital prices typically fell toward the lower end of the prediction range, or closer to the price that sellers (having no knowledge of the bonus) were willing to take than to the absolute highest price that producers would pay, losing all of the bonus through high price. With the average capital price and the price range, about 18 percent of the bonus was bid away in the form of higher prices or implicit taxes. Table 6 also indicates that 57 percent of the labor and capital choices are within one unit of predictions, which, given the numerous input combinations for any given output, is also highly encouraging. The results in the table suggest that the isoquants were deliberate choices shaped by relatively efficient market forces.

In summary, there is support for our conjecture that producers in the tax benefit experiments moved to lower isoquants. This in turn helps explain why our original hypothesis of a capital purchase increase (with the introduction of accelerated tax depreciation) was not supported; at these lower isoquants, optimal units of capital purchased (on average, 12) are identical to optimal units of capital purchased in the benchmark experiments. The fact that our implicit tax hypothesis was supported lends support to the power of competitive markets. That is, the higher exogenous tax subsidy was partly bid away to capital sellers through the operation of supply and demand.

IV. Conclusions and Extensions

In this study, the level of capital asset investment was not affected by a tax policy similar to ITC or accelerated depreciation. This result is inconsistent with extant neoclassical theory and the expectations of policymakers. The level of investment was lower than expected in part because the prices of capital assets were bid up. That is, there was shifting of the tax benefits backward to the capital suppliers in the form of "implicit taxes." As previously discussed, findings were not consistent with theory insofar as producers in the tax benefit experiments did not move to the highest isoquant possible, and hence did not purchase enough total inputs or produce enough output. One conjecture here is risk aversion on the part of producers; that is, they were uncertain whether they could sell all of their output at a sufficiently high price to cover costs.[15] By underpurchasing inputs and underproducing, producers erred on the side of conservatism. Alternatively the observed underproduction could be explained by subjects' lack of rational expectations. In experimental market settings with probabilistic payoffs, Lundholm (1991) and Plott and Sunder (1988) observed that rational expectations limited the equilibrating abilities of a single market. In our study, producers are required to project output markets rationally and equate them with two input markets. In fact, the setting is more similar to the earliest models of rational expectations—cobweb models of agriculture in which farmers plant on the basis of expected prices in the next season. In the cobweb model, adaptive expectations result in a gradual attainment of market clearing. In our setting, this would imply a more gradual equilibrating process as producers learn the output demand curve.

---

[15] An inspection of unsold inventories in table 3 and 4 tends to corroborate this. In the last three periods, average unsold inventories (across all experiments) were 12.5 percent of production (7.6 percent if Experiment 3 is excluded). Verbal post-experimental interrogatories of producers from Experiments 3 and 6 indicated that they had sufficient time to sell all inventories but experienced occasional difficulty selling a last unit or two at prices they considered sufficient.

A final potential explanation relates to incentives. Although producers should have been motivated to produce at the higher isoquant (they each would have recieved, on average, an extra dollar per period over their chosen isoquant levels), each output buyer would have recieved, on average, only 2 to 3 cents for the additional units purchased at the higher isoquant. A number of experiments have indicated subjects' willingness to trade for such modest profits, but it may have been that the cost of trading exceeded the benefit—a failure of dominance cannot be dismissed with respect to these marginal units.

While it might be argued that either risk aversion or lack of rational expectations could explain the observed behavior in the tax benefit markets, the persuasive force of such arguments is reduced by the random assignment of subjects across the two experimental conditions. Neither underproduction nor underpurchasing were generally observed in the benchmark markets. This suggests that the move to a lower-than-predicted isoquant in response to a tax benefit may be properly attributed to the experimental treatment. This unexpected result calls for additional research.

Tax policies regarding capital (depreciation and ITC) are intended to stimulate fixed asset investment. However, the results in the current study suggest that these policies will have an attenuated effect because of tax shifting.[16] Furthermore, some of the tax benefits will accrue to factor suppliers and consumers. Such shifting is consistent with the results reported by Kachelmeier et al. (1991), who found costs could be shifted, in varying degrees, forward to consumers in a single-market setting.

The study also has implications for an emerging line of research examining multiple markets with production. In the only published study in this area (to our knowledge), Goodfellow and Plott (1990) examined one input and one output market.[17] A set of subjects bought inputs and converted them into outputs using a production function specified as $x_i = 7y_i - (y_i^2/2)$, where x and y are units of ouput and input, respectively. Prices and quantities converged toward CE predictions, but never actually achieved CE with stability. As in our experiments, prices and quantities varied between experiments, and there were occasional unsold inventories. Nonetheless, efficiency (defined as the actual surplus of consumers and producers taken as a percentage of the maximum possible) was high, and the authors concluded that the markets behaved "as if" the competitive assumption were satisfied. Our results are similar and, by applying the same qualitative criterion, appear encouraging in the sense that markets behaved "as if" the competitive assumptions were satisfied.

---

[16] We must note that, in the real world, long run shifting price effects may be mitigated if additional capital suppliers can enter the market. The time needed for an industry to reach a long-run equilibrium depends primarily on barriers to entry for supplies (e.g., technical, scale, capital, and market barriers) and the shape of the average cost curve. The length of this lag will vary between industries. For example, an increase in the demand for automobiles generated by a tax subsidy will result in a slow supply response (if any) because of high (technological and cost) barriers to entry into the automobile industry. Likewise, the effect of a tax policy (e.g., a 10 percent investment tax credit) may be too small to even affect long-run supply, given barriers to entry. Hammermeth (1980, 756) demonstrates that labor supply has a lagged response to tax changes across industries and suggests that such effects may occur for capital supply (plant and equipment) as well. In the extreme instance, constantly changing tax policy (as is presently observed in the United States) may completely thwart long-run adjustments. Similarly, the existence of uncertainty regarding future tax policy may have this same effect.

[17] In a study employing multiple markets (although not a production economy per se), Kachelmeier et al. (1992) examine the relative implicit tax effects of three forms of consumption tax (a sales tax, a gross receipts tax, and a value-added tax) in a setting with wholesalers, retailers, and consumers. Unlike the results reported here, their results generally support theoretical predictions.

Appendix A
Intermediate Steps in Solving for Equilibria

Intermediate Steps for Deriving Equation (9) in Text

Substituting $w_1 x_1 = (1-\beta) w_2 x_2$ for x_1 in the production function yields (at the individual producer's level):

$$w_2 x_2 = \left(\frac{y}{3}\right)^{1.25} \bullet (w_1 w_2)^{0.5} \bullet n^{0.25}. \tag{A1}$$

The total cost function is:

$$C(y, w_1, w_2) = w_1 x_1 + (1-\beta) w_2 x_2 = 2(1-\beta) w_2 x_2, \tag{A2}$$

which is then substituted into equation (A1) to yield equation (9).

Intermediate Steps for Deriving Equation (11) in Text

Since there is perfect competition, firms will produce until $MC = p$, thus:

$$y^{0.25} = \frac{2}{5} p \bullet \left(\frac{3^{1.25}}{n^{0.25}}\right) \bullet [(1-\beta) w_1 w_2]^{-0.5}, \tag{A3}$$

so that firm output supply is:

$$y^s(p) = \frac{16}{625} \bullet \left(\frac{3^5}{n}\right) \bullet p^4 \bullet [(1-\beta) w_1 w_2]^{-2}. \tag{A4}$$

We then derive aggregate output market clearing $y^D(p) = n \bullet y^s(p)$, as follows:

$$\frac{\alpha}{P} = \left(\frac{16}{625}\right) \bullet 3^5 \bullet p^4 \bullet [(1-\beta) w_1 w_2]^{-2}, \tag{A5}$$

so that,

$$P = \left(\frac{625\alpha}{16}\right)^{0.2} \bullet \frac{1}{3} [(1-\beta) w_1 w_2]^{0.4}. \tag{A6}$$

Thus, aggregate output market-clearing demand is:

$$y^D = \frac{\alpha}{P} = \frac{3\alpha}{\left(\frac{625\alpha}{16}\right)^{0.2} \bullet [(1-\beta) w_1 w_2]^{0.4}}, \tag{A7}$$

which is simplified in equation (11).

Intermediate Steps for Deriving Equations (12) and (13) in Text

The firm's input demand function for x_2 is:

$$x_2 = 3^{-1.25} \bullet \frac{y^{1.25}}{n} \bullet \left(\frac{w_1}{(1-\beta) w_2}\right)^{0.5} \bullet n^{0.25}. \tag{A8}$$

We then substitute aggregate demand equation (A7) into the firm's input demand function for x_2, which yields:

$$x_2^D=\left(\frac{2\alpha}{5}\right)\bullet\frac{1}{(1-\beta)w_2}\bullet\frac{1}{n}. \tag{A9}$$

Input market clearing for X_2 implies

$$X_2^S(w_2)=n\bullet x_2^D[(1-\beta)w_2] \tag{A10}$$

at the firm level. At the aggregàte level,

$$\frac{N}{n}\left(\frac{2\alpha}{5}\right)[(1-\beta)w_2]^{-1}=bw_2,$$

which is then substituted into previous equations to yield equations (12) and (13). Similarly, aggregate X_1, demand is:

$$X_1^D=\frac{2\alpha}{5w_1}. \tag{A11}$$

Aggregate Supply Function for Producers

Combining previous definitions yields:

$$Y^S=1.56\frac{p^4}{[(1-\beta)w_1w_2]^2}. \tag{A12}$$

To determine integer-form supply, the least-cost combinations of inputs for each Y (in terms of each possible combination of labor and capital, at prices determined from the sellers' supply functions) were substituted into the denominator of equation (A12) to solve for p at each integer level of Y (note that p was rounded up to the nearest integer).

Appendix B

Near-Equilibrium Input and Output Combinations for Producers

There are two combinations for the benchmark condition (no tax benefit), as shown below.

	Combination 1	*Combination 2*
Labor		
Quantity	24	28
Price Range	5–8	6–7
Total Cost	156	182
Capital		
Quantity	12	12
Price Range	16–22	16–22
Total Cost	171	171
Output		
Quantity	28	32
Price Range	11–17	14–15
Total Cost	392	464

As can be seen, Combination 2 is cost-minimizing since the total costs for both factors are roughly equivalent and closer to the market-clearing prices and quantities (figs. 2 through 4). Note that only symmetric solutions for inputs (i.e., each buyer chooses identical inputs) are considered here since theory is at individual producer level. Only interior solutions are considered (i.e., price and quantity solutions equal to or less than CE quantities). See figures 2 through 4 for integer-form prices for supply and demand. Total cost of capital is average market price less bonus.

The three combinations for the tax benefit condition are as follows:

	Combination 1	2	3
Labor			
Quantity	24	28	28
Price Range	5–8	6–7	6–7
Total Cost	156	182	182
Capital			
Quantity	12	12	16
Price Range	16–33	16–33	22–25
Total Cost	147	147	188
Output			
Quantity	28	32	36
Price Range	9–17	11–15	13
Total Cost	364	416	468

As can be seen, Combination 3 is the cost-minimizing solution closest to competitive equilibrium prices and quantities, although we note that Combination 1 is also roughly cost-minimizing. See the preceding discussion for other assumptions.

References

Bischoff, C. 1971. Business investment in the 1970s: A comparison of models. In *Brookings Papers on Economic Activity*, edited by A. Okun and G. Perry. Washington, D.C.: Brookings Institution, 13–58.

Chirinko, R. 1986. Business investment and tax policy: A perspective on existing models and empirical results. *National Tax Journal* 39 (June): 137–53.

——, and R. Eisner. 1982. The effect of tax parameters in the investment equations in macroeconometric models. In *Economic Activities and Finance*, edited by M. Blume et al. Cambridge, MA: Ballinger.

——, and ——. 1983. Tax policy and investment in major U.S. macroeconomic econometric models. *Journal of Public Economics* 20 (March): 139–66.

Coen, R. 1971. The effect of cash flow on the speed of adjustment. In *Tax Incentives and Capital Spending*, edited by G. Fromm. Washington, D.C.: Brookings Institution, 131–96.

Davis, J., and C. Swenson. 1988. The role of experimental economics in tax policy research. *Journal of the American Taxation Association* 10 (Fall): 40–59.

Goodfellow, J., and C. Plott. 1990. An experimental examination of the simultaneous determination of input prices and output prices. *Southern Economic Journal* 56 (April): 969–83.

Hall, A., and D. Jorgenson. 1967. Tax policy and investment behavior. *American Economic Review* 57 (June): 391–414.

Hammermeth, D. 1980. Factor market dynamics and the incidence of taxes and subsidies. *Quarterly Journal of Economics* 95 (December): 751–64.

Harberger, A. 1962. The incidence of the corporation income tax. *Journal of Political Economy* 70 (June): 215–40.

Henderson, J., and R. Quandt. 1980. *Microeconomic Theory: A Mathematical Approach*. New York: McGraw-Hill.

Kachelmeier, S., S. Limberg, and M. Schadewald. 1991. A laboratory market examination of the consumer price response to information about producers' costs and profits. *The Accounting Review* 66 (October) 694–717.

——, ——, and ——. 1992. Tax revenue and incidence under alternative consumption tax regimes. Unpublished manuscript, University of Texas at Austin.

Lundholm, R. 1991. What affects the efficiency of a market? Some answers from the laboratory. *The Accounting Review* 66 (July): 486–515.

Plott, C. 1991. A computerized laboratory market system and research systems for the multiple double auction. Social Science Working paper No. 783, California Institute of Technology, Pasadena.

———, and S. Sunder. 1988. Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica* 56 (September): 1085–1118.

Scholes, M., and M. Wolfson. 1992. *Taxes and Business Strategy: A Planning Approach*. Englewood Cliffs, NJ: Prentice-Hall.

Smith, V. 1982. Microeconomic systems as an experimental science. *American Economic Review* 72 (December): 923–55.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 515–533

The FASB's Policy of Extended Adoption for New Standards: An Examination of FAS No. 87

Russell Langer
San Francisco State University
Baruch Lev
The University of California at Berkeley

SYNOPSIS AND INTRODUCTION: A *multi-year* adoption period now appears to be the norm for new accounting standards issued by the Financial Accounting Standards Board (FASB). For example, FAS No. 52, which pertained to foreign currency, was issued in December 1981 and became effective in 1983, a three-year adoption period. FAS No. 71, which concerned regulation, was issued in December 1982, but became effective in 1984. FAS No. 87 on pensions was issued in December 1985 and became effective in 1987, but a key provision of this statement—the recognition of a "minimum liability"—became effective only in 1989, thereby allowing a five-year adoption period. FAS No. 96 on income taxes was issued in December 1987, but amendments under FAS Nos. 100 and 103 extended its adoption to 1990 and then to 1992, a six-year adoption period (ultimately, FAS No. 109 substantially changed FAS No. 96 and its successors).

Although several early statements (e.g., FAS Nos. 2 and 8) issued late in the calendar year allowed adoption over that year and the following one, practically all FASB statements issued in the 1970s became effective on a uniform date close to their issuance.[1] Thus, it appears that the FASB changed its adoption policy in the 1980s. The Board's main justification for an extended adoption period is to alleviate firms' implementation costs,

---

[1] Two notable exceptions are FAS Nos. 13 and 19. FAS No. 13 became effective for new leases January 1, 1977, but its retroactive application to existing leases was deferred until 1981. This extension was overruled by the SEC, which required retroactive application in 1978. FAS No. 19 concerned oil and gas producers and was issued in December 1977 but not required until 1979. This statement was vetoed by the SEC before it became effective.

The helpful comments and suggestions of Eli Amir, Mary Barth, Somnath Das, Mark DeFond, Bala Dharan, Amir Guttman, Wayne Landsman, Toshiyuki Shibano, Andrew Rose, Brett Trueman, as well as those of the three reviewers are gratefully acknowledged.

Submitted June 1991.
Accepted September 1992.

particularly the costs of renegotiating agreements with lenders and suppliers (see, e.g., FAS No. 52, par. 147-48, and FAS No. 87 par. 259-60).

Interestingly, no justification was offered for the extended adoption period of FAS No. 96 (Income Taxes), perhaps an indication that a three-year adoption period had become a norm. However, when a fourth year was added under FAS No. 100, the following rationale was given (FASB 1988, par. 8):

> The Board believes that the disadvantages to preparers from not having a deferral of the effective date outweigh the disadvantages to users from a one-year delay in the required adoption of Statement 96, including diversity in financial reporting from the continued application of Opinion 11 by some enterprises.

Given the obvious costs imposed by an extended adoption period on financial statement users, because of reduced cross-company comparability, the FASB's policy warrants scrutiny.[2] This is the objective of the current study which focuses on FAS No. 87 (FASB 1985a) and the related FAS No. 88 (FASB 1985b). Specifically, we consider various possible managerial motives for choosing the timing of adoption of FAS No. 87 within the allowed period, and classify these motives as involving either *compliance costs* (the FASB's express justification for an extended adoption period) or *investor perceptions* (managers' attempts to change investor expectations). We then identify the adoption timing motives that are consistent with the data derived from samples of early (1986) and late (1987) adopters. The FASB's case for extending adoption periods will obviously be supported if compliance costs figure predominantly in firms' adoption-timing decisions.

Of the eight proxies for adoption-timing motives examined by us, only one—increasing reported earnings—consistently discriminates between early and late adopters. This holds for interyear (1986 vs. 1987) as well as intrayear adoptions (first three quarters of 1986 vs. fourth quarter). Of the compliance-cost motives examined, company size and the number of outstanding loans were associated with the adoption-timing decision in some cases. Overall, our analysis does *not* provide compelling support for the FASB's cost-reduction justification for a multiyear adoption period for FAS No. 87.

Key Words: *FASB policy, Compliance costs, Earnings management, Pensions.*

Data Availability: *A list of the sample firms is available upon request.*

---

[2] It is difficult to reliably estimate these private and social costs, but they are not negligible. One source of such costs was highlighted in our detailed examination of firms' disclosures concerning FAS No. 87 which revealed numerous cases of inadequate and fuzzy disclosures; these complicate the cross-sectional adjustment of financial statements to a uniform basis.

I. Factors Affecting the Timing of FAS No. 87 Adoption

THE managerial motives that are conjectured to affect the adoption timing of FAS No. 87 are classified into two categories: (1) compliance cost, cost factors that will induce some firms to adopt FAS No. 87 early and others to postpone adoption to save costs, and (2) investor perception motives, which concern managers' attempts to change the preceptions of investors about the value of the company. This distinction between adoption-timing motives is not always clear-cut, but was chosen since the FASB's justification for extending the adoption period of new standards is primarily based on compliance cost. Distinguishing *a priori* between cost and perception motives and identifying the specific motives consistent with the data will, therefore, allow us to evaluate the justification for the FASB's policy of extended adoption.

Compliance Costs

Recontracting Costs. The FASB frequently justifies an extended adoption period as intended to allow firms time to recontract obligations whose covenants might be violated by the accounting change.[3] A comprehensive examination of this issue as it relates to accounting changes requires both a detailed analysis of the sample firms' individual obligations and related covenants as well as an assessment of the costs (e.g., higher interest rates) associated with recontracting. This is obviously a difficult task, and the cost assessment is particularly intractable. As a first pass at the issue as it relates to FAS 87, we focus on the debt/equity ratio, a frequently used proxy for the restrictiveness of loan covenants.

Recent research indicates that the debt/equity ratio is a reasonably good proxy. For example, Duke and Hunt (1990, 61) report that "all seven versions of the debt/equity proxy are positively related to the existence and tightness of retained earnings restrictions. All seven proxies are also related to the existence of working capital and net tangible asset restrictions. The three restrictions captured by the [debt/equity] proxy variables include the two most prevalent types of accounting-based restrictions. . . ." Also, Beneish and Press (1992) report that the median debt-to-total-asset ratio of their sample firms that violated loan agreements was significantly larger than that of a control sample of nonviolators, which indicates that the debt/equity ratio reflects the restrictiveness of loan covenants reasonably well.

Since adoption of FAS No. 87 had a different effect on earnings and equity values across firms, our expectation regarding the relation between the debt/equity ratio and adoption timing should be conditional on this effect. Accordingly, we predict that firms whose earnings and equity *increased* with adoption of FAS Nos. 87 and 88 will tend to adopt early when their debt/equity ratio is high, since such adoption will alleviate loan constraints to some extent. In contrast, companies for which FAS 87 adoption *decreased* earnings and equity will tend to adopt late when their debt/equity ratio is high, because these companies presumably need, according to the FASB, additional time to renegotiate loan agreements. Our prediction therefore is:

---

[3] For example, "The Board decided to allow an additional two years [for the 'minimum liability' requirement] . . . because of concerns expressed by some respondents that some employers would have to arrange to renegotiate or obtain waivers of provisions of some legal contracts" (FAS No. 87, par. 260).

P1: Firms with a positive effect of adoption on earnings and a relatively high debt/equity ratio will tend to adopt early, while firms with a negative adoption effect and a high debt/equity ratio will adopt late.[4]

Since the debt/equity ratio is not a complete measure of the cost of loan renegotiation, we examined an additional proxy, the number of outstanding loans and bond issues.[5] *Ceteris paribus*, the larger the number of loans (and lenders), the more costly and time-consuming the process of renegotiation will be. Here too, the expectation is conditioned by the effect of adoption on earnings and equity. Our expectation is:

P2: Firms with a positive effect of adoption on earnings and a relatively large number of loans will tend to adopt early, while firms with a negative adoption effect and a larger number of loans will tend to adopt late.

Direct Costs of Adopting FAS NO. 87. The renegotiation and default costs depicted above relate to the consequences of adopting FAS No. 87. In addition, one should consider the direct out-of-pocket costs incurred in collecting the data needed to implement the statement. Most of the new data relate to the computation of the projected benefit obligation (PBO) and the actuarial method on which it is based, namely, the "projected unit credit method." Some firms adopted this actuarial method after the FASB indicated its preference for it in the Preliminary Views issued in 1982.[6] Accordingly, firms that were already using this method prior to the issuance of FAS No. 87 should have encountered fewer difficulties in coping with the complexity of the standard than firms that had not adopted this actuarial method. This leads to the following prediction:

P3: Firms already using the projected unit credit method will adopt FAS No. 87 earlier than firms using other actuarial methods.

Additionally, the adoption costs of large companies are believed to be lower than those of small companies since the former typically generate more information for internal purposes and use outside consultants (e.g., actuaries, appraisers) more frequently.[7] Assuming that firms with relatively low compliance cost are motivated to adopt accounting standards early,[8] we predict:

P4: Large firms will tend to adopt FAS No. 87 earlier than small firms.

Managing Investor Perceptions

Undervaluation of the Pension Surplus (Overfunding). Empirical studies suggest that, although the pension surplus (pension assets minus pension liabilities) was recognized by investors as an asset prior to FAS No. 87 (i.e., the surplus was reflected in the market value of the firm), investors tended to discount the value of the surplus (see, e.g., Landsman and Ohlson 1990). A reasonable explanation for such investor discounting is that the pension surplus reported in footnotes prior to FAS No. 87 was based on the difference between the fair value of pension assets and the pension liability valued by the accumulated benefit obligation (ABO) method. The ABO value is generally *lower* than the PBO value because the former abstracts from expected increases in future compensation levels. If, as is widely believed (see, e.g., Barth 1991; Ippolito 1986), the PBO reflects the pension liability more appropriately than the ABO (e.g., the PBO is more consistent with the economic pension liability and with the "going concern" assumption), then investors were justified in discounting the reported pension surplus.

---

[4] Although firms with low debt/equity ratios are not excluded from the sample, this prediction is not particularly applicable to them, nor is the next prediction, P2, for firms with a small number of loans.

[5] This variable was suggested by a reviewer.

[6] See Ghicas (1990).

[7] According to FAS No. 87 (par. 209), "Some respondents argued that smaller employers would have a more difficult time than other employers with the initial application of this Statement, in part because advisors involved in pension accounting may put a higher priority on the needs of larger employers. The Board agreed that the transition provisions of this Statement, which allow an extra two years before application is required for certain smaller employers. . . ."

[8] One of the most frequently mentioned reasons for early adoption was that the information required (e.g., on the pension liability under PBO) was already on hand, and managers saw no reason to wait until the standard would be generally required (from phone conversations with early adopting managers).

PBO values, however, were not required to be reported prior to FAS No. 87; hence, investors' discount of the pension surplus was probably based on some industry- or economy-wide average expectation of PBO values. Accordingly, managers of firms with low PBO values relative to ABOs would be motivated to adopt FAS No. 87 early to report these low PBOs to investors and allow them to correct the undervaluation. We predict:

P5: Early adopters of FAS No. 87 are characterized by lower values of PBO relative to ABO than late adopters.

Managing Reported Earnings. Not much is known about managers' motives to "manage" earnings, the means used for such management, and its capital market consequences. However, evidence suggesting the prevalence of earnings management abounds (see, e.g., Dharan and Lev 1993; Hand 1990; McNichols and Wilson 1988). Absent specific theoretical or empirical guidelines to form an "earnings management" hypothesis, we chose two simple conjectures: (1) managers are motivated to report higher earnings whenever they can do so, and (2) managers of firms that experience an annual earnings decrease will attempt to mask this decrease by adopting an income-increasing accounting change.[9] Accordingly, we predict:

P6: Firms that adopted FAS No. 87 early had increased their reported earnings more than late adopters.

P7: Firms choosing early adoption had a lower (less favorable) earnings change *without* the adoption than did late adopters.

Disguising a Too-High Discount Rate. Prior to FAS No. 87, firms enjoyed considerable leeway in selecting the discount rate for computing the present value of the pension obligation. Managers that wished to report a lower-than-actual pension liability, and thereby a lower underfunded position (or a higher pension surplus), could select, within reasonable bounds, a high discount rate. Suspecting such behavior, the FASB narrowed considerably the latitude of discount rates with the issuance of FAS No. 87, generally prescribing the use of the "settlement rate."[10] Thus, managers that used a higher-than-appropriate discount rate prior to the adoption of FAS No. 87 would have had to report a significantly lower discount rate upon adoption, possibly calling investors' attention to the discount rate manipulation. By postponing adoption, such managers could gradually decrease the discount rate (over fiscal years 1985 and 1986), so that the drop in the discount rate upon adoption in 1987 would hardly be noticeable. Accordingly, our prediction is:

---

[9] This conjecture is empirically supported by Dharan and Lev (1993): the majority of their sample firms that increased reported earnings by an accounting change would have reported an earnings decrease without the change.

P8: Early adopters of FAS No. 87 are characterized by a low pension discount rate in the pre-adoption period relative to late adopters.

The validity of the outlined adoption-timing motives, classified into compliance cost and perception-changing categories, will now be examined on samples of early and late adopters of FAS No. 87. Recall that several of the preceding arguments about adoption-timing motives were industry-specific (e.g., the pension surplus undervaluation might be related to investors' use of industry-wide expectations of PBO values). Accordingly, in the following empirical tests, measures that might be industry-specific—the debt/equity ratio (P1), the PBO relative to ABO (P5), and the earning change (P7)—are defined relative to their industry averages.

II. Sample Selection and Variable Definition

The samples of early and late FAS 87 adopters were selected from the 600 firms listed in the 1987 edition of *Accounting Trends and Techniques*. We restricted the samples to firms listed on the New York Stock Exchange and having December fiscal year-ends.[11] Annual reports and 10-K forms for the years 1985, 1986, and 1987 were examined to collect pension-related and other data, primarily from the pension footnotes. Additional information was obtained from the COMPUSTAT and CRSP tapes. Firms not available on COMPUTSTAT and CRSP were deleted from the samples. The sample selection is summarized as follows:

	Accounting Trends and Techniques' Sample	*NYSE and December Fiscal Firms*	*Final Sample (COMPUSTAT and CRSP)*
1985 Adopters	18	12	–
1986 Adopters	267	181	151
1987 Adopters	275*	93	74
TOTAL	560	286	225

*Includes 1988 adoptions by firms with non-December year-ends.

Because of the small number of 1985 adoptions, we restricted the final sample to 1986 and 1987 adopters, although the 12 firms listed for 1985 are discussed in section IV.

---

[10] The settlement rate is the rate implicit in annuity contracts that could be used to settle the pension obligation. Although not a unique rate, it is based on external evidence and is more objective than the expected rate of return used to discount the ABO under FAS No. 36.

[11] The first restriction was mandated by the sources we used for accounting and stock price data. The second was made because firms with non-December year-ends had a different adoption period, 1986–1988.

The 225 firms in the final sample are distributed over 34 two-digit SIC industries. Six industries are represented by more than ten companies each (i.e., 16 firms in industry No. 26, paper and plastics; 15 firms in No. 29, petroleum; 14 firms each in Nos. 33, steel and metals, and 34, fabricated metals; 12 firms in No. 36, electrical equipment, appliances, and electronic components; and 23 firms in No. 37, motor vehicles and aircraft). Financial firms (banks, insurance companies, etc.) are not represented in our sample.

The year of adoption was defined as the year in which a firm adopted FAS No. 87 for its U.S. pension plans. Because adoption of FAS No. 88 required the prior adoption of FAS No. 87, some of the variables described below reflect the *combined* effect of these standards on earnings.

Of the final sample of 151 early (1986) adopters, 131 firms disclosed the effect of adoption on either the pension expense, net income, and/or earnings per share (EPS). When only the effect on pension expense was disclosed, we estimated the effect on EPS by using the firm's effective tax rate for 1986 and the average number of common shares used for the EPS computation. It is noteworthy that the voluntary adoption of FAS No. 87 in 1986 had a non-negative effect on the earnings of all but two sample firms. Even for these two firms, the combined effect of FAS Nos. 87 and 88 was substantially positive (see fn. 15).

The variables representing the hypothesized motives for adoption timing were defined as follows:

P1 (*recontracting costs*), the industry-adjusted book value of the debt/equity ratio (total debt divided by total equity, common and preferred) as of 12/31/85 was used as a proxy for the restrictiveness of loan covenants.[12]

P2 (*recontracting costs*), the number of outstanding loans and bond issues at the end of 1986 as reported by *Moody's Industrial Manuals*.

P3 (*actuarial method*), a dummy variable to distinguish between firms that adopted the projected unit credit method prior to 1986 and those that adopted the method later.

P4 (*size*), firm size measured as the log of total assets on 12/31/85 as a proxy for some direct compliance costs related to FAS No. 87.

P5 (*undervaluation of the pension surplus*), the firm's PBO/ABO ratio, adjusted for the industry.[13] Because most companies did not disclose the PBO value until the end of the year of adoption, we used the PBO/ABO ratio at the end of 1986 for the early adopters and the PBO/ABO ratio at the end of 1987 for the late adopters. We examined the stability of the PBO/ABO ratio from the end of 1985 to the end of 1987 on a sample

---

[12] The industry adjustment of the debt/equity ratio was:

$$\frac{debt/equity^F - debt/equity^I}{debt/equity^I},$$

where $debt/equity^F$ is the firm's ratio, and $debt/equity^I$ is the average ratio for the two-digit SIC industry to which the firm belongs. We computed both weighted and unweighted industry averages. The industry adjustment for measures P5 (PBO/ABO) and P7 (earnings change) is identical to that of the debt/equity ratio.

[13] An empirical examination of the variable:

$$\frac{\text{Pension assets minus PBO}}{\text{Pension assets minus ABO}}$$

yielded similar empirical findings.

of 50 firms that disclosed PBO values as of December 31, 1985. The mean change of the PBO/ABO ratio over these two years was only −1.7 percent, an indication that it is reasonable to assume that the late adopters PBO/ABO ratio for 1986 was close to the 1987 reported value.[14]

P6 (effect of adoption on earnings) is measured as the effect of adopting FAS Nos. 87 and 88 on EPS, deflated by share price at the beginning of 1986 (to enhance cross-sectional comparability). Because adoption of FAS No. 88 requires adoption of FAS No. 87, we considered the combined effect as a single variable.

Under FAS No. 88, the asset reversion that occurs when a defined benefit pension plan is terminated and replaced with another plan is recognized immediately as a gain. Previously, these reversions had been recorded as deferred credits and amortized over ten years. The retroactive application of FAS No. 88 to a reversion in a prior year is recognized as a cumulative effect of a change in an accounting principle. In our sample, 14 firms had reversions in 1986, 13 of which chose early adoption of FAS Nos. 87 and 88. Nine firms in our sample had reversions prior to 1986, seven of which elected early adoption of FAS Nos. 87 and 88. The EPS variable in this study is income from continuing operations and therefore the "adoption impact" variable (P6) includes the 1986 reversions but not the pre-1986 reversions. It is worth noting that the two sample firms (out of 151) that reported a negative effect on earnings from early adoption of FAS No. 87 had relatively large positive cumulative effects from applying FAS No. 88 to pre-1986 reversions.[15] Thus, *no* sample firm's bottom line was negatively affected by the early adoption of FAS Nos. 87 and 88.

As mentioned in section I, the FASB's extended adoption policy probably imposes nontrivial costs on financial statement users since adjusting cross-sectionally for different accounting techniques resulting from different adoption timings is not straightforward. For example, the effect of FAS No. 87 on earnings was disclosed in footnotes by 131 of the 151 early adopters in our sample. Of the 20 nondisclosures, most were explained by "immateriality," which might lead an analyst to believe that no adjustment is required.[16] However, the immateriality of some of these nondisclosures appears questionable, as illustrated by Lockheed's case. Lockheed did not disclose the adoption effect on 1986 earnings, stating that it was "impracticable" to trace it through the work-in-process inventory and immaterial because the pension plan was at the ERISA funding limit. However, given that Lockheed funded $56 million in 1986 (information derived from the pension footnote), which probably would have been reported as pension expense under APB No. 8,[17] and that Lockheed reported in 1986 a *negative* $40 million pension expense (see data below), the effect of adopting FAS No. 87 in 1986 was to *increase* pretax earnings by $96 million. This amounts to a 79-cent increase in

---

[14] This small change in PBO/ABO is not surprising. The PBO/ABO ratio is directly related to the projected rate of salary increase assumed by a firm in calculating the PBO. *Accounting Trends and Techniques 1988* reports virtually no change in the distribution of projected rates of salary increases used by the firms in their surveys of 1986 and 1987.

[15] These firms are Fairchild Industries and Reynolds Metals. For Fairchild, the dollar effect of FAS No. 87 on EPS was (0.005), and the effect of FAS No. 88 was 3.17. The corresponding figures for Reynolds were (0.68) and 3.09.

[16] FAS No. 87 did not specifically require disclosure of the adoption impact.

[17] Only 20 percent of the 1986 adopters in our sample had prepaid (accrued) pension cost as of January 1, 1986, which suggests that the pension expense equaled funding under APB No. 8 for the remaining 80 percent of the sample.

EPS and is approximately *ten times* Lockheed's reported increase in EPS for 1986 (from $6.10 to $6.18, see data below). McDonnell-Douglas, operating in the same industry and having a pension plan with a virtually identical funded status, apparently had no difficulty determining the effect of FAS No. 87 adoption ($2.71 per share).[18] A comparison of the pertinent data for Lockheed and McDonnell-Douglas is presented below. All amounts, except EPS, are in millions of dollars.

	Lockheed (1986)	*McDonnell-Douglas (1987)*
Dec. 31 ABO	$3,084	$2,971
Dec. 31 PBO	3,536	3,328
Dec. 31 pension assets at fair value	4,526	4,474
Dec. 31 transitional assets	1,001	1,520
Pension expense (income):		
Previous year (APB No. 8)	159	165
Current year (FAS No. 87)	(40)	(103)
Effect of adopting FAS No. 87:		
On pension expense	Not reported	Not reported
On net income	Not reported	109
On EPS	Not reported	2.71
EPS:		
Previous Year	6.10	6.86
Current year	6.18	7.75

Questionable "immateriality" cases, such as Lockheed's, demonstrate the difficulties analysts encounter in cross-sectionally adjusting financial data for different adoption timings.[19] This motivated us to *estimate* the effect of FAS No. 87 on earnings for the 20 firms that adopted the standard in 1986 without disclosing its effect (see appendix). The tests reported below were conducted by using these estimates alternatively with a zero effect (i.e., accepting the company's immateriality claim) for the nonreporters. For the late (1987) adopters, we estimate what the earnings effect would have been had they adopted FAS No. 87 in 1986, using the estimation method described in the appendix. Alternatively, we used the actual 1987 effect disclosed under the assumption that the effect would have been similar had they adopted in 1986.[20]

---

[18] McDonnell-Douglas' decision to delay adoption until 1987 seems surprising, given the favorable effect on EPS. However, the aggregate cost actuarial method used by McDonnell-Douglas on its government contract bids had to be used for financial reporting also because of a conformity requirement in Cost Accounting Standard No. 412. McDonnell-Douglas was either unable or unwilling to discontinue using this method until December 1986, at which time it adopted the "entry-age normal method" that was not subject to a conformity requirement. McDonnell-Douglas was then able to adopt FAS No. 87, with its projected unit credit method, in 1987. In contrast, Lockheed was already using the entry-age normal method and thus could adopt FAS No. 87 in 1986.

[19] As observed by Kidder, Peabody Equity Research (1988, 1), "The market has generally not looked beyond company earnings per share to differentiate between those with operating earnings improvements and those resulting from lower pension costs."

[20] The correlation between the effect reported by late adopters for 1987 and our estimates of the effect had they adopted FAS No. 87 in 1986 is very high: Pearson correlation coefficient = 0.95.

Table 1
Sample Means for Early (1986) and Late (1987) Adopters
FAS Nos. 87 and 88

Variable	*Early Adopters* (n=151)	*Expected Relation*	*Late Adopters* (n=74)	p-*value*
P1: Debt/Equity	1.431 (1.279)		1.635 (2.518)	0.4342
P2: Number of Loans	10.199 (7.720)		6.922 (5.233)	0.0014
P3: Actuarial Method (dummy)	0.218 (0.414)	>	0.231 (0.424)	0.8240
P4: Total Assets ($ Millions)	5,476 (10,785)	>	3,077 (4,607)	0.0816
P5: PBO/ABO	1.227 (0.184)	<	1.211 (0.172)	0.5517
P6: Effect on EPS (Divided by Price)	0.0077[a] (0.0146)	>	0.0010[a] (0.0080)	0.0002
	0.0085[b] (0.0144)	>	0.0015[b] (0.0099)	0.0003
P7: Pre-Adoption Earnings Change (Divided by Price)	0.0034[a] (0.1197)	<	0.0074 (0.1630)	0.8325
	0.0026[b] (0.1194)	<	0.0074 (0.1630)	0.7980
P8: Pre-Adoption Discount Rate	8.295 (0.852)	<	8.411 (1.074)	0.3704

Note: Standard deviations are in parentheses. The *p*-value of chi-square is based on a univariate Logit analysis.

[a] The effect of the standards for nondisclosing early adopters is assumed to be zero; for late adopters, it is the actual effect reported for 1987.

[b] The effect for nondisclosing early adopters and for late adopters is estimated (see appendix).

P7 (*pre-adoption earnings change*) is measured as the 1986 EPS from continuing operations minus the 1985 EPS from continuing operations, deflated by share price at the beginning of 1986. For 1986 adopters, the 1986 EPS was adjusted to its pre-adoption level by subtracting the tax-adjusted effects of FAS No. 87 and 88. Since we used EPS from continuing operations, calculation of this variable was not affected by the cumulative effect of FAS No. 88 on pre-1986 terminations.

P8 (*disguising the discount rate*) is measured as the interest rate used by sample firms in 1985 to discount the ABO, as disclosed in the 1985 annual reports.

III. Motives for Adoption Timing: Findings

Summary statistics for the variables representing the adoption motives are given in table 1. The differences between the mean values for the early and late adopters are in the predicted direction for four of the eight variables: total assets (P4), adoption effect

(P6), earnings change (P7), and pre-adoption discount rate (P8). On a univariate basis, the differences between the means are statistically significant (at the 10 percent level) for two variables only: total assets and adoption effect on EPOS. Both are larger for early adopters. None of the "wrong-direction" variables are statistically significant. The variables debt/equity (P1) and number of loans (P2) do not lend themselves to a straightforward prediction of the difference in means, since the prediction is conditional on whether the effect of adoption on earnings is positive or negative. The data, however, indicate that early adopters had, on average, a significantly larger number of loans than late adopters.

The univariate statistical significance (or lack of significance) ignores the interrelationships among the independent variables. To account for these, we performed a multivariate Logit analysis, combining the samples of early and late adopters and considering all the adoption motives simultaneously. The Logit estimates of four models are presented in table 2. Models 1, 2, and 3 are based on our estimates of the effect of adoption on 1986 EPS for all the late (1987) adopters, as well as for the 20 early adopters that did not disclose an effect (see appendix). To check the sensitivity of our findings to these estimates, model 4 is based on the *actual* effect on EPS disclosed by the late adopters in 1987, as well as on using a zero effect for the 20 early adopters that did not disclose an effect. Given the relatively high correlation between firm size (P4) and number of loans (P2), we also examined the effect of including both variables (models 3 and 4) as well as only one variable (model 2) in the Logit runs. The four models in table 2 are thus intended to examine the sensitivity of our findings to different estimates and model configurations.

The only variable that emerges as statistically significant in all four models is P6, the effect of adoption on EPS. The number of loans (P2) is the only other variable significant in model 1 (the model without the slope dummies), an indication that early adopters tended to have a significantly larger number of loans than late adopters. Recall, however, that the prediction regarding number of loans was conditional on the effect of adoption on earnings: companies with a positive effect on earnings and a large number of loans are motivated to adopt FAS No. 87 early, while those with a negative effect will adopt late. This prediction is examined and confirmed in model 3. The variable for number of loans is run with a slope dummy assigned the value of 1 for a positive or no adoption effect on earnings, and 0 for a negative effect. The positive and statistically significant coefficient for the dummy-times-loans variable (P2A) indicates that companies with a large number of loans *and* a positive effect of adoption on earnings tended to adopt FAS No. 87 early; the negative and significant coefficient for the loans variables (P2) indicates that companies with a negative effect of adoption *and* a large number of loans tended to adopt FAS No. 87 late (i.e., in 1987).[21] It should be noted, however, that the number of loans is highly correlated with firm size (Pearson correlation coefficient $=0.55$). Omitting the number of loans from the Logit analysis in model 2 yields a significant coefficient for P4, firm size ($p=0.08$). None of the other variables examined, including the debt/equity ratio with a slope dummy (P1A), yielded statistically significant coefficients in any of the models run.

---

[21] We also ran Logit models with interaction variables rather than slope dummies (i.e., variable P6 times Loans, and P6 times Debt/Equity). The only statistically significant variable in these models was P6 times Loans ($p=0.01$). However, the p-value for variable P6 alone was only 0.13, probably because of the very high correlation between this variable and the interaction, P6 times Loans variable (Pearson coefficient $=0.89$).

Table 2
Logit Estimates of Hypothesized Adoption-Timing Motives

Independent Variable	*1*	*2*	*3*	*4*
Intercept	0.371 (0.87)	−2.200 (0.29)	1.188 (0.64)	−1.233 (0.62)
P1: Debt/Equity[a]	−0.127 (0.31)	−0.148 (0.42)	−0.039 (0.84)	−0.239 (0.76)
P1A: Dummy[b]×Debt/Equity	-	0.197 (0.49)	−0.092 (0.76)	0.325 (0.70)
P2: Number of Loans	0.085 (0.02)	-	−0.434 (0.04)	−0.134 (0.37)
P2A: Dummy[b]×No. of Loans	-	-	0.553 (0.01)	0.227 (0.13)
P3: Actuarial Method	−0.317 (0.40)	−0.331 (0.37)	−0.299 (0.46)	−0.390 (0.33)
P4: Log Total Assets	0.034 (0.81)	0.198 (0.08)	−0.070 (0.65)	0.044 (0.76)
P5: PBO/ABO[a]	−0.411 (0.34)	−0.495 (0.24)	−0.591 (0.29)	1.724 (0.11)
P6: Adoption Effect on EPS	102.40 (0.00)	111.80 (0.00)	44.18 (0.08)	117.40 (0.01)
P7: Pre-Adoption Change in EPS[a]	1.385 (0.20)	1.529 (0.13)	0.715 (0.57)	1.267 (0.31)
P8: Pre-Adoption Discount Rate	−0.138 (0.42)	−0.028 (0.87)	−0.032 (0.87)	0.049 (0.80)
Chi-Square	27.11 (0.00)	22.18 (0.00)	52.08 (0.00)	35.45 (0.00)

Note: The dependent variable is 1 for early adopters (1986), 0 for late adopters (1987); p-values are in parentheses. In models 1 through 3, the effect of adoption on EPS is estimated as described in the appendix for late adopters and for early adopters that did not disclose an effect. In model 4, the effect is as actually disclosed by late adopters in 1987; unreported effects are assumed to be zero for both early and late adopters.

[a] This variable is industry-adjusted (value-weighted); see footnote 12. Logit estimates are qualitatively identical when the industry average is equally weighted.

[b] Dummy = 0 for firms when the effect of adoption on EPS was negative, 1 when the effect was zero or positive.

Except for the correlation between firm size and number of loans, multicollinearity does not appear to pose a serious problem in our analysis. None of the correlations between the variables used in the Logit analysis (except the slope dummies with the underlying variables) exceeds 0.20. Also, using the 1987 *reported* effect on EPS for the late adopters—model 4—does not yield materially different results from those of the other three models where the late adopters' earnings effect (had they adopted in 1986) was estimated.

In summary, the effect of adoption on reported EPS is the only variable that is consistently significant in the Logit analysis.[22] The number of loans and firm size also appear to be related to the adoption-timing decision, although the high correlation between the two variables makes it difficult to disentangle their individual effects on the adoption decision. The importance of the effect on EPS for the timing of adoption is also reported by Scott (1991), who examined various political and agency variables affecting the adoption of FAS No. 87. Scott's finding that debt variables (e.g., leverage) appear unrelated to the adoption decision is also consistent with ours.[23]

IV. Some Characteristics of 1985 and 1987 Adopters

The finding that increasing reported earnings was the major motive for the early adoption of FAS No. 87 raises the question whether this also motivated the 12 very early adopters in 1985 that were not included in the Logit analysis. It also raises the question of why late (1987) adopters that could have increased their 1986 earnings significantly by adopting FAS No. 87 did not do so. An individual examination of such cases might shed further light on managerial adoption-timing motives.

Of the 12 firms in our original sample that adopted No. FAS 87 (and FAS No. 88 if applicable) in 1985, nine disclosed the effect of adoption on earnings: the effect was positive in all nine cases, with per share amounts ranging from 3 to 54 cents (average EPS relative to share price: 0.005). The remaining three firms that did not disclose the earnings effect had positive transitional assets, an indication that the effect was not negative. Eight of the 12 firms had pre-adoption earnings decreases relative to 1984. Five were already using the suggested actuarial method for FAS No. 87 before adoption. Thus, the overall characteristics of the 1985 adopters, particularly the positive effect on reported earnings, are similar to those of the early (1986) adopters.

As to the late (1987) adopters that could have increased substantially their 1986 reported earnings by an early adoption, we selected for individual examination the companies that reported for 1987 that the effect of adoption on EPS (deflated by price) was larger than 0.0077, the average effect on EPS for the 1986 adopters (see table 1). The following six companies met this criterion. Cleveland Cliffs and Oak Industries reported large losses in 1986 and would still have reported losses in 1986 even if they had adopted FAS No. 87. These firms' earnings might have been below the lower bound of management compensation plans, and adoption would not have benefitted managers. Their decision to postpone adoption is also consistent with the "big bath" phenomenon: Interlake Corporation terminated a major portion of its pension plans in 1987, so adoption of FAS No. 87 in 1986 for plans that were soon to be terminated seems pointless. As mentioned earlier, the actuarial method of McDonnell-Douglas precluded early adoption of FAS No. 87. Both Grumman and Northrop stated that the savings from adopting FAS No. 87 would have to be passed along to their customers—primarily the U.S. government to which both firms made over 90 percent of their sales under cost-plus contracts.[24] Apparently, these companies did not feel rushed to pass the savings to this major client. Overall, the late (1987) adopters with a large, positive effect on earnings appear to have had good reasons to postpone adoption.

---

[22] We also ran the Logit models reported in table 2 with *firm-specific* values for variable P1, P5, and P7, rather than the *industry-adjusted* variables reported. The strong significance of P6, the variable for the effect of adoption on earnings, is confirmed with the firm-specific values.

[23] Scott's study differs from ours in several important respects. The objective of our study is to evaluate the FASB's extended adoption policy; Scott, however, endeavors to identify political and agency variables affecting the adoption decision. Furthermore, we examine both inter- (1986 vs. 1987) and intrayear (1986, see Section V) motives for adoption, while Scott's study is restricted to the former. Also, Scott's sample is restricted to companies that increased their reported earnings by adopting FAS No. 87; our sample includes firms that decreased their 1987 reported earnings with adoption. Scott examines variables that we do not (and vice versa), such as the regulation status of the company and the existence of a bonus plan tied to income. Finally, our study includes the effect of FAS No. 88, while Scott's does not.

V. Choosing the Quarter of Adoption

Early adopters of FAS No. 87 (in 1985 and 1986) risked foregoing an opportunity to report higher earnings subsequent to the adoption year, for the following reason. The transitional asset (i.e., the difference between the value of the plan assets and the PBO) that is credited to earnings over a certain period (generally 15 years) is based on the *beginning-of-the-year* overfunded balance. Gains or losses on pension assets because of changes in security prices, or on the value of the PBO because of variation of the actuarial assumptions (e.g., the discount rate) *subsequent* to the beginning of the adoption year could be credited to income only if greater than a "corridor": 10 percent of PBO or plan assets, whichever is greater. And the credit is limited to the excess of this gain or loss over the corridor. Early adoption, therefore, could have prevented a firm from reporting higher earnings in the future if the overfunded balance increased during the adoption year, and could have denied managers higher compensation. Consider the following example:

AT&T's transitional asset on 1/1/86 was $7.7 billion. Since AT&T adopted FAS No. 87 in the first quarter of 1986 and amortized the overfunded balance over 16 years, it increased 1986 pretax earnings by $484 million. During 1986, AT&T's overfunded balance increased by $1.3 billion, because of increases in market prices of securities. Had AT&T postponed the adoption by one year, it could have increased reported pretax income by about $80 million per year for the next 16 years ($1.3 billion divided by 16). This income increase was "lost" since the 1986 increase in the overfunded balance was smaller than the corridor of 10 percent of the plan assets ($26.6 billion).

A rational policy for a firm considering adoption of FAS No. 87 in 1986 seemed to be to wait until the end of 1986 (or early 1987), observe the actual change in the overfunded balance for 1986 and then decide whether to adopt FAS No. 87 for 1986 or postpone adoption to 1987. However, as indicated by the data below, close to 60 percent of our sample of 1986 adopters chose to adopt the new statement in one of the first three quarters of the year, thereby denying themselves this choice.[25]

Quarter	*Number of Adoptions*
1st	59
2nd	17
3rd	8
4th	57
Unknown	10

---

[24] Unlike McDonnell-Douglas, Grumman and Northrop were not using the aggregate cost method, so it is not clear to us why they would have to pass the savings along. Perhaps they were concerned with the effect on their next round of contracts.

[25] One of the few firms to adopt in the third quarter was Goodyear Tire and Rubber, which "coincidentally" terminated a pension plan during the third quarter and recognized a gain of $1.39 per share under FAS No. 88. Goodyear's reported EPS for 1986 including this gain was $0.94.

Table 3
Logit Estimates of Intra-1986 Adoption Motives

Independent Variable	*Coefficient*	p-*Value*
Intercept	−2.1426	0.4503
P1: Debt/Equity[a]	0.0440	0.8484
P2: Number of Loans	−0.0252	0.4000
P3: Actuarial Method	0.1957	0.6795
P4: Log Total Assets	0.3646	0.0356
P5: PBO/ABO[a]	1.6929	0.2168
P6: Effect on EPS[b]	32.7327	0.0893
P7: Earnings Change (Annual)[a,b]	1.5357	0.3266
P8: Pre-Adoption Discount Rate	−0.3293	0.1509
P9: Potential Gain	206.0	0.0504
Model Chi-Square	14.175	0.1162

Note: The dependent variable is 1 for first- to third-quarter adopters ($n=84$), 0 for fourth-quarter adopters ($n=57$).

[a] Industry-adjusted.

[b] The effect of adoption for nondisclosing firms is assumed to be zero.

What was the reason for the impatience of the early 1986 adopters? The decision-relevant variable for the intra-1986 adoption is the increase in *subsequent* years' reported EPS resulting from postponing adoption until 1987. This potential reporting gain in subsequent EPS is computed as follows:

$$\text{Potential Gain} = \left[\frac{(\text{Gain}_{86}/15)\times(1-\text{tax rate})}{\text{number of shares outstanding}}\right] / \text{beginning stock price.}$$

Gain_{86} is the change between 1/1/86 and 12/31/86 in the balance of pension assets minus PBO. Assuming a 15-year amortization, 1/15 of that gain will be credited to earnings starting in 1987. Recall that only fourth-quarter 1986 adopters had the information on this potential gain at the time of adoption. We therefore expect this variable to be smaller, on average, for fourth-quarter adopters than for earlier adopters, since the former presumably adopted FAS No. 87 in 1986 only if the observed potential gain was small (less than the corridor), while earlier adopters either used some average expectation of the gain in the adoption decision or ignored this consideration altogether.

To examine this expectation and the motives for the intra-1986 adoption of FAS No. 87, we ran an intra-1986 Logit analysis giving the dependent variable a value of 1 for adopters in the first three quarters of 1986 and 0 for fourth-quarter adopters. Logit estimates are reported in table 3. The potential gain variable, P9, is positive and statistically significant, as expected. First- to third-quarter adopters in 1986 indeed gave up larger future increases in reported earnings than fourth-quarter adopters. The reasons for this can be inferred by the statistically significant coefficients in table 3.

The effect of adoption on EPS (P6) is positive and statistically significant, indicating that first to third quarter adopters enjoyed a higher earnings increase than fourth-quarter adopters. The only other significant coefficient in table 3 is that of firm size (P4): first- to third-quarter adopters were larger than fourth-quarter adopters. However, because adoption costs, proxied by firm size, do not play an important role in the intrayear adoption decision (as contrasted with the interyear decision), the effect on reported earnings is left as the major discriminatory variable between early and late 1986 adopters. Thus, it appears that in order to report higher EPS *earlier* in 1986, some firms were willing to forgo higher reported earnings (and perhaps some managerial compensation) in subsequent years, in case the pension surplus increased during 1986. So strong is the temptation to increase reported EPS as soon as possible.

VI. Discussion

A multiyear adoption period for new accounting standards now appears to be the FASB's norm. Although this policy obviously increases users' costs through decreased cross-sectional comparability of financial information, the benefits of the policy in lowering preparers' costs (compliance and contract renegotiation) are less obvious. One way of assessing these benefits is to identify managers' motives in choosing the timing of adoption within the period allowed by the FASB. If these motives are strongly related to adoption costs, then the FASB's case for an extended adoption period is supported.

The analysis reported fails to provide a strong support for cost reduction through extended adoption period. Of the various proxies for compliance and contract renegotiation cost examined, only the number of loans and company size are found to differ significantly between early and late adopters. Yet, even if the costs of adopting FAS No. 87 were relatively higher for small firms or for those with a large number of loans, it is not obvious how much could be saved by *postponing* the adoption decision by one or two years.

Early adoption of FAS No. 87 appears to be strongly related to the increase in reported earnings: all of the 151 early adopters in our sample increased their reported earnings with the adoption of FAS Nos. 87 and 88. Given the strong showing of the variable for the effect of adoption on EPS, it is tempting to conclude that *earnings management* was at work here. This conclusion is supported by other studies that examined early versus late adoptions. For example, Pincus and Wasley (1991, 5) examined close to 3,700 cases of mandatory accounting changes related to 16 FASB standards and concluded that "the earnings impact of mandated accounting changes was positive and more so for early adopters relative to late adopters. . . . These results support the conjecture that their early adoptions of mandated accounting changes were used as a tool for earnings management."[26] Also recall that adoption of FAS No. 87 is needed prior to adoption of FAS No. 88 and that a number of firms in our sample of early adopters (14 out of 151) indeed terminated their pension plans in 1986. Such termination further increases earnings by allowing firms an immediate recognition of the prepaid pension costs, rather than amortizing it over several years.[27] Thus, the immediate effect on reported earnings appears to be an important consideration in the adoption timing decision of new accounting standards.

---

[26] Ayres (1986) also reports that early adoption of FAS No. 52 had a larger effect on earnings than late adoption.

The costs of a multiyear adoption period imposed on financial statement users, the absence of compelling evidence supporting the FASB contention that an extended adoption period is needed to alleviate compliance costs, and the existence of pervasive evidence that increasing earnings motivate early adoptions, should, in our opinion, cause the FASB to reconsider its multiyear adoption policy.[28] However, it may be that the FASB's real reason for granting an extended adoption period is not economic but political, fending off the strong antagonism of major preparer groups to practically every new substantive accounting standard. An evaluation of such a political regulatory motive is beyond the boundaries of this study.

Finally, what explains managers' evident passion for increasing reported earnings by early adoption of accounting standards? Affecting management compensation, says the conventional wisdom.[29] We feel, however, that there is more to the earnings enhancement story than increasing compensation. The compensation argument cannot explain, for example, our findings concerning the intra-1986 quarterly adoption decisions. Obviously, our knowledge of managerial motives in disclosure decisions is incomplete, and puzzles, like those reported here, should be pursued to enhance this knowledge.

Appendix

Estimating the 1986 Earnings Effect of FAS No. 87

We estimate the effect of FAS No. 87 on 1986 earnings for the early (1986) adopters that did not disclose this information and the late (1987) adopters *had* they adopted in 1986. We compared the pension expense computation under FAS No. 87 and that under APB Opinion No. 8. The main differences are:

1. *Actuarial method.* The projected unit credit method, required by FAS No. 87, is a benefit allocation method that reflects projected future salaries but not future years of service. In contrast, APB Opinion No. 8 permitted several actuarial methods, and most firms used cost allocation methods that reflect both projected future salaries and future years of service. Accordingly, adoption of the projected unit credit method tends to decrease the "normal cost" (or service cost) component of the pension expense. Although "normal cost" was not separately disclosed under APB Opinion No. 8, it was a major component of the pension expense. Accordingly, we use the pension expense under APB Opinion No. 8 as a surrogate for normal cost in the estimation of the effect of FAS No. 87 on earnings.
2. *Amortization.* The transitional obligation (asset) existing at the date of initial adoption of FAS No. 87 represents unrecognized prior service costs from plan amendments and unrecognized gains (losses) prior to adoption of FAS No. 87. Amortization of the transitional balance under FAS No. 87 (typically over 15 years) is generally faster than the amortization of prior service cost and gain (loss) under APB Opinion No. 8. We thus expect the effect of FAS No. 87 on pension expense to be positively (negatively) correlated with the transitional obligation (asset) at the date of adoption.
3. *Interest and expected earnings.* Under FAS No. 87, interest cost is used for the *PBO* computation and expected earnings for the plan assets. Under APB Opinion No. 8, interest cost (income) was calculated on the accrued (prepaid) pension cost recognized on the balance sheet. At the date of FAS No. 87 adoption, the transitional obligation (asset) is calculated as:

$$PBO-[ASSETS+ACCRUED(PREPAID)PENSION\ COST].$$

---

[27] Clinch and Magliolo (1991, 16) comment that such pension settlements "appear to be an effective (and low cost) means for firms with overfunded pension plans to manage their reported income without any tax effects."

[28] Beneish and Press (1991) provide evidence that firms' cost of default on loan covenants is, on average, small, further suggesting that the FASB's concern with such costs (a major justification for the extended adoption period) might be exaggerated.

[29] Indeed, Scott (1991) finds that the existence of an income-based bonus plan influences management's FAS No. 87 adoption choice.

We thus expect the effect of adoption on pension expense to be positively (negatively) correlated with the transitional obligation (asset) at the date of adoption, not only because of the amortization element (point 2 above) but also because of the manner of calculating interest cost and expected earnings (point 3).

According to the above arguments, the pension expense under FAS No. 87 should be related to three variables presented in regression form:

$$IPX/TA=a+b_0PENX/TA+b_1(D\times PENX/TA)+b_2TRAN/TA+\epsilon,$$

where:

IPX = the effect of FAS No. 87 on pension expense in the year of adoption,
TA = total company assets at the beginning of the year of adoption,[30]
$PENX$ = pension expense in the year prior to adoption,
$TRAN$ = transitional asset (obligation) at the beginning of the adoption year,[31] and
D = dummy variable: 1 for firms that used the FAS No. 87 mandated actuarial method prior to 1986, and 0 for firms that used other actuarial methods.

The above regression was run cross-sectionally by OLS on: (1) the 131 early (1986) adopters that disclosed the effect of adoption on the 1986 pension expense, and (2) the 42 late (1987) adopters that disclosed the effect of adoption on the 1987 pension expense. The regression estimates (*t*-values in parentheses) are as follows:

	1986 Adopters	1987 Adopters
Intercept (a)	−0.002	0.001
	(−3.889)	(2.07)
Pension expense (b_0)	−0.188	−0.286
	−4.733)	(−6.865)
Dummy × Pension expense (b_1)	0.079	−0.039
	(0.996)	(−0.489)
Transitional asset (b_2)	−0.080	−0.100
	(−14.072)	(−16.901)
Adjusted R^2	0.61	0.90
Observations	131	42

The coefficients for the pension expense and transitional asset (obligation) variables are highly significant and have the expected signs, and the R^2s are reasonably high, lending credence to our estimation method.

To apply the estimation method to the nondisclosing 1986 adopters, we multiplied the coefficients from the 1986 adopters regression (given above) by the $PENX_{85}$ and $TRAN_{85}$ of the nondisclosing firms to estimate the effect on pension expense. We then applied the effective tax rate and divided by the number of outstanding shares to estimate the effect on EPS. To apply the estimation method to the 1987 adopters, we first had to estimate the 1985 transitional asset (obligation) for these firms, as follows:

$$TRAN_{85}=ASSETS_{85}-\left[\frac{PBO_{87}}{ABO_{87}}\times ABO_{85}\right].$$ [32]

We then multiplied the coefficients from the 1987 adopters regression by the $PENX_{85}$ and $TRAN_{85}$ variables for those firms, applied the effective tax rate, and divided by the number of shares.

---

[30] We deflated the variable to alleviate heteroscedasticity.

[31] Since most firms did not disclose the beginning-of-the-year balance, we assumed 15-year amortization and multiplied the ending balance of the transitional asset (obligation) by 15/14.

[32] $ASSETS_{85}$ = pension plan asset value at the end of 1985. See footnote 14 in the text for the justification of using the 1987 values of ABO and PBO.

References

Ayres, F. 1986. Characteristics of firms electing early adoption of SFAS No. 52. *Journal of Accounting & Economics* 8(2) 143–58.

Barth, M. 1991. Relative measurement errors among alternative pension asset and liability measures. *The Accounting Review* 66 (July): 433–63.

Beneish, M., and E. Press. 1991. Consequences and costs of technical violation of debt covenants. Working paper, Fuqua School of Business, Duke University, Durham, North Carolina.

——, and ——. 1992. Technical default and stock prices. Working paper, Fuqua School of Business, Duke University, Durham, North Carolina.

Clinch, G., and J. Magliolo. 1991. Executive compensation effects of accounting-oriented discretionary transactions in commercial banks. Working paper, Wharton School, Philadelphia.

Dharan, B., and B. Lev. 1993. The valuation consequences of accounting changes: A multi-year examination. *Journal of Accounting, Auditing & Finance* (Forthcoming).

Duke, J., and H. Hunt. 1990. An empirical examination of debt covenant restrictions and accounting-related debt proxies. *Journal of Accounting & Economics* 12 (January): 45–63.

Financial Accounting Standards Board. 1985a. *Statement of Financial Accounting Standards No. 87: Employers' Accounting for Pensions*. Stamford, CT: FASB.

——. 1985b. *Statement of Financial Accounting Standards No. 88: Employers' Accounting for Settlements and Curtailments of Defined Benefit Pension Plans and for Termination Benefits*. Stamford, CT: FASB.

——. 1988. *Statement of Financial Accounting Standards No. 100: Accounting for Income Taxes Deferral of the Effective Date of FASB Statement No. 96*. Stamford, CT: FASB.

Ghicas, D. C. 1990. Determinants of actuarial cost method changes for pension accounting and funding. *The Accounting Review* 65 (April): 384–405.

Hand, J. 1990. A test of the extended functional fixation hypothesis. *The Accounting Review* 65 (October): 739–63.

Ippolito, R. 1986. *Pensions, Economics and Public Policy*. Dow Jones-Irwin.

Kidder, Peabody Equity Research. 1988. Pension Expense. (August).

Landsman, W., and J. Ohlson. 1990. Evaluation of market efficiency for supplementary accounting disclosures: The case of pension assets and liabilities. *Contemporary Accounting Research* 7 (Fall): 185–98.

McNichols, M., and P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

Pincus, M., and C. Wasley. 1991. Time-series and cross-sectional characteristics of accounting changes. Working paper, Washington University, St. Louis, MO.

Scott, T. 1991. Pension disclosures under SFAS No. 87: Theory and evidence. *Contemporary Accounting Research* 8 (Fall): 62–81.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 534–561

Disclosure Policy and Competition: Cournot vs. Bertrand

Masako N. Darrough
University of California—Davis

SYNOPSIS AND INTRODUCTION: Disclosure of financial information is an essential ingredient of a well-functioning capital market. However, public disclosure of information can affect a disclosing firm negatively if market participants make strategic use of the information to their advantage. In the presence of such a "proprietary cost," a firm has to trade off the positive and negative effects of disclosure.

In an oligopolistic environment, disclosure causes rival firms to respond. The response depends on the nature of competition and private information. Some firms benefit by hiding, and others by sharing, information. If firms do not disclose information voluntarily, mandating disclosures will force firms to disclose information that they wish hidden. Mandating has no incremental effect if firms would have voluntarily disclosed the information. To promote more efficient (welfare-maximizing) disclosure policies, it is essential to understand how firms would behave in the absence of mandatory disclosure requirements. The purpose of this article is to analyze that behavior.

A two-stage model of a duopoly is formulated to analyze firms' incentives to disclose private information. The incentives depend on whether firms are engaged in Cournot or Bertrand competition, and whether the private information is about demand or cost. Both *ex ante* incentives to precommit to a disclosure policy and *ex post* incentives to disclose voluntarily are examined.

Ex ante, firms would not commit to disclosure in Cournot/demand and Bertrand/cost cases. Although Cournot duopolists would not commit to disclosure of information about demand, both firms and consumers might be better off *if disclosure were enforced* by regulatory agencies such as the Securities Exchange Commission (SEC) or the Financial Accounting Standards Board (FASB). Firms would commit to share information in the cases

I am grateful to comments by Dov Fried, Paolo Fulghieri, Jack Hughes (associate editor), Alison Kirby, Shinji Kobayashi, Murgie Krishnan, Ram Ramakrishnan, Neal Stoughton, and the participants at the seminars at New York University, University of California- Davis, University of Maryland, Purdue University, the Far Eastern Econometric Socity, and the American Accounting Association meetings. Financial support from the Graduate School of Business, Columbia University, is acknowledged.

Submitted November 1991.
Accepted December 1992.

of Cournot/cost and Bertrand/demand. However, firms' incentives diverge *ex post* because the benefit of disclosure depends on the realized value of the signal. When the existence of private information is suspected, but not disclosed, nondisclosure is attributed to the type of signal that is better undisclosed. Thus, in equilibrium, it is difficult for firms to hide information successfully. In the Cournot/demand case, virtually all values of private information would be disclosed. In contrast, in Bertrand/cost, disclosure would seldom be observed when products are good substitutes.

The model developed in this article identifies the environments in which mandatory disclosure rules are most effective: (1) when firms have the incentive to precommit to nondisclosure and (2) when voluntary disclosure is least likely in the absence of precommitment.

Key Words: *Disclosure, Cournot and Bertrand competition,* ex ante *and* ex post *incentives.*

I. Background and Overview

FIRMS release financial reports to various stakeholders to provide timely and relevant information that is useful for investment decisions, monitoring and rewarding performance, and writing contracts. Since the demand for financial information comes from various sources (shareholders, creditors, employees, suppliers, government agencies, etc.), it is possible that the disclosure of particular information has differential effects on these parties. For example, detailed disclosure about new products conveys information about the future prospects of a firm to its shareholders. But it might also reveal strategic information to competitors, thereby reducing the disclosing firm's competitive advantage. The disclosure in this case involves both positive and negative effects on the welfare of shareholders of the firm. The negative effect is often referred to as a "proprietary cost."[1] In the presence of such a cost, a firm has to trade off the positive against the negative effects of disclosure. Revealing information to competitors, however, does not always reduce the disclosing firm's future profit. In fact, in some situations, firms are better off sharing information so as to coordinate actions to their mutual advantage. The consequences of disclosure depend on the specific type of competition firms are engaged in and the type of private information firms have.

Since there are conflicting incentives for disclosure, it is not at all clear whether firms will *voluntarily* disclose all relevant information. Anticipating potentially conflicting incentives, the users of financial information will try to infer the underlying information that is withheld. In equilibrium, firms are successful in withholding information only in limited cases. Mandating disclosures through regulatory agencies such as the SEC or the FASB will force firms to disclose the type of informaton that firms wish hidden. In such a case, mandating has a real effect on the workings of the market, with potentially different effects on the stakeholders. In other situations, mandating might have no incremental effect because firms would have voluntarily disclosed the information anyway. It is important for regulatory agencies to sort out the disclosure incentives of firms in order to promote more efficient disclosure policies that are consistent with the goals of the agencies.[2] In particular, it is crucial to understand how firms would behave in the absence of mandatory disclosure requirements. Only then is it possible to determine the welfare implications of mandating various disclosure requirements. The purpose of this article is to provide the necessary analysis.

---

[1] Foster (1986) uses the term "competitive disadvantage costs."

The particular setting considered is a two-stage, noncooperative game of duopoly with private information. Two firms, which are engaged in either Cournot or Bertrand competition, decide on disclosure in the first stage, and quantity (or price) in the second stage. Firms might commit themselves in advance to a particular disclosure policy before they receive their private information on demand or cost. Alternatively, firms might choose their disclosure strategy after the receipt of private signals. In making decisions, firms take into account the strategic effect of their decisions on their rival firms. If firms find it beneficial to share private information, they are likely to precommit themselves to such a disclosure *ex ante*. (This commitment may be coordinated and enforced by the FASB or the SEC.) Once firms receive signals, however, some firms might find that withholding the information would have been better. The precommitted disclosure policy need not be sequentially rational. Whether firms want to precommit themselves to a particular disclosure policy *ex ante*, or whether they prefer to disclose private information voluntarily *ex post* depends critically on the type of competition they are engaged in and the type of private information they receive.

If firms are willing to disclose information on a voluntary basis, mandating disclosure is redundant. If voluntary disclosure is not forthcoming, however, mandatory disclosure can significantly affect the welfare of various stakeholders. For example, both Cournot duopolists and consumers might be better off if disclosure of information on demand is enforced, even though, *ex ante*, the duopolists would not commit to disclosure voluntarily.

Such a clear-cut Pareto improvement does not exist for the case of Bertrand competition with cost information. *Ex post*, firms make disclosure decisions on the basis of the realized values of the signals they receive. Their decisions also depend on how nondisclosure is interpreted by the market. For example, if firms are expected to possess private information, nondisclosure will be attributed to their having information that is better undisclosed, rather than to lack of private information. In such an environment, it is difficult to hide information. Even in environments in which firms would have made prior commitments to nondisclosure, if they failed to precommit, they are more likely than not to disclose the information. Virtually all values of information would be disclosed when firms compete Cournot-style with demand uncertainty. Disclosure, however, is less likely in the Bertrand/cost combination. In fact, disclosure is likely to be rarely observed, when two products are very good substitutes.

The contribution of this article is best understood by placing it in the literature on disclosure incentives. An important result in Grossman (1981) and Milgrom (1981) is that full disclosure is the unique equilibrium when firms are privately informed and are concerned with financial market valuation. The impetus for disclosure comes from the desire of better-type firms to communicate with the financial market. Since a lack of disclosure indicates that a firm is of a worse type, the firm is forced to disclose to differentiate itself from even worse firms. Thus, the whole process unravels so that full disclosure obtains.

---

[2] Although we place the SEC and the FASB in the same category, it is quite possible that their goals differ in many respects. We abstract from these issues in this article.

Typical assumptions in these full disclosure models are that (1) it is common knowledge that firms have private information; (2) if firms disclose information, they do so truthfully; and (3) firms are concerned with financial market valuation. Without these assumptions, full disclosure might not obtain. If the market does not know whether firms have private information, then a lack of disclosure will not necessarily be attributed to bad news, but possibly to lack of private information (Jung and Kwon 1988). If certain types of information are not verifiable, truthful disclosure is not always credible, which results in partial disclosure of only "certifiable" information (Okuno-Fujiwara et al. 1990). Nondisclosure might also prevail because of the existence of proprietary costs. Verrecchia (1983) shows that in the presence of (exogenously imposed) proprietary costs, a manager withholds information when it falls below some threshold value. The threshold, furthermore, decreases with the quality of information (Verrecchia 1990b).

Proprietary costs have been endogenized in the context of an entry game (Darrough and Stoughton 1990; Dye 1986; Feltham and Xie 1992; Wagenhofer 1990) in which an incumbent considers the effect of disclosure on both financial and product markets.[3] An incumbent with favorable information about demand has conflicting incentives, since good news might raise its valuation in the financial market but might also trigger the entry of competitors into the market. Conversely, an incumbent with unfavorable information might suffer downgrading in the financial market but might succeed in deterring entry by a disclosure.

An interesting implication of the analysis in Darrough and Stoughton (1990) is that competition through a threat of entry encourages voluntary disclosure. The strongest incentive for disclosure to discourage entry comes from an incumbent with unfavorable information, since entry takes place only if the prospect is favorable. Darrough and Stoughton conclude that when entry deterrence is important, the unique equilibrium is one of full disclosure.[4] By construction, entry deterrence is given disproportionate importance, in part to highlight the notion of proprietary costs. The question of how competition in general affects disclosure behavior remains unresolved. If one takes the viewpoint that Bertrand is more competitive than Cournot competition (for each level of substitutability), or associates the degree of substitutability with competition, then the finding in this article is consistent with the conjecture that the less competitive an industry, the more likely disclosure is to prevail (Verrecchia 1983).

An extensive literature also concerns disclosure in the context of information sharing through the voluntary (precommitment) mechanism of trade associations.[5] Most of the models have analyzed settings of duopoly or oligopoly to establish that the type of competition, whether Cournot or Bertrand, makes a significant difference in choosing, *ex ante*, whether to share information on demand or cost.[6]

---

[3] A similar analysis was made in the context of duopoly by Dontoh (1990) and Gertner et al. (1988).

[4] Verrecchia (1990a) correctly points out that in some sense an entry game "exaggerates the usefulness of 'bad news' to discourage market entrants. . . ."

[5] See, e.g., Clarke (1983), Fried (1984), Gal-Or (1985, 1986), Novshek and Sonnenschein (1982), Shapiro (1986), and Vives (1984).

[6] In a recent article, Vives (1990) focuses on monopolistic competition, which prevents any single firm from influencing the aggregate market.

Although there are similarities between trade associations and regulatory agencies, there are also important differences. Trade associations collect and disseminate information on behalf of their members. Membership is voluntary, and information sharing can be exclusionary (i.e., can exclude nonmembers from obtaining information). In contrast, the SEC or the FASB has enforcement power. Compliance is mandatory, and most disclosure requirements are of a public nature. In addition, the constituencies these regulatory agencies serve are varied. For these reasons, the SEC and the FASB can choose a policy that enhances desired social objectives, even at the expense of some stakeholders. For example, line-of-business reporting can be interpreted as a requirement of accurate disclosure mandated by the FASB. Feltham et al. (1992) interpret aggregate-profit reporting as noisy disclosure (or less information sharing) and line-of-business reporting as noiseless disclosure (full information sharing). In Hughes and Kao (1991), capitalizing R&D spending reveals firms' costs, while expensing does not.

In this article, I analyze disclosure incentives with the particular goal of identifying how the types of competition and private information influence incentives for voluntary disclosure. This is necessary because the incentives for, and effects of, disclosure critically depend on the specific configuration of competition and private information. A market with two firms already in place (duopoly) is better suited for this purpose than an entry game, since duopoly reduces the effect of a rival's behavior. Thus, bad news would not play a disproportionate role in strategic behavior. By focusing on the effect of disclosure on the product market, one is able to classify when a disclosure involves a proprietary cost and when it results in beneficial outcomes. In particular, the findings are that mandatory disclosure rules are most effective: (1) when firms have the incentive to precommit to nondisclosure and (2) when voluntary disclosure is least likely in the absence of precommitment.

The remainder of this article is organized as follows. In the next section, I present a two-stage model of duopoly under both Cournot and Bertrand competition with private information on demand or cost. The incentive to disclose *ex ante* is analyzed by having firms decide on a disclosure policy before they receive a signal on the uncertain parameter. In the *ex post* setting of section III, firms that have not precommitted to a specific disclosure policy decide on disclosure after the receipt of a signal. Different types of firms will have different incentives, depending on how nondisclosure will be perceived by the competitor, and I delineate how disclosure incentives are affected by the type of competition and the nature of the information. It is shown that in the Cournot/demand case, virtually all signals will be disclosed, whereas in the Bertrand/cost case, firms can succeed in hiding information more frequently. The final section presents a brief summary and concluding remarks. All proofs are presented in the appendix.

II. *Ex Ante* Setting

The *ex ante* setting has been analyzed extensively in the context of information sharing via trade associations that collect and disseminate information on behalf of their members.[7] Since members give instruction as to what type of information is to be collected and disclosed, they in effect choose, and commit themselves to, a disclosure policy ex *ante*. In this analysis, I focus on disclosure policies that are mandated and enforced by regulatory agencies. Firms lobby for or against various disclosure principles, such as generally accepted accounting principles (GAAP) or disclosure requirements for exchange listing, some of which allow more leeway with respect to "quality" or "the degree of aggregation." Disclosure with noise, however, need not be fraudulent; it might simply be imprecise or not very informative. Since I focus on information on *future* demand and costs, it is quite plausible that disclosure does not provide a completely accurate representation of the private signals received. It is assumed that disclosure is therefore truthful in the sense of not containing fraudulent information, but that, depending on the level of noise chosen, information content or quality may vary. Further, I assume that whatever is disclosed has already been verified by auditors or is easily verifiable at a negligible cost.[8]

---

[7] Gal-Or (1985) and Kirby (1988) study incentives for information sharing in a Cournot oligopoly with demand uncertainty, whereas Shapiro (1986) focuses on cost uncertainty. Both Cournot and Bertrand settings are analyzed by Vives (1984) with demand uncertainty, and by Gal-Or (1986) with cost uncertainty.

Benchmark Under Certainty

Before uncertainty about demand or cost is introduced, it is worthwhile to review a standard model of duopoly under Cournot or Bertrand competition. Assume a linear inverse demand function of the form for firm i, i=1, 2,[9]

$$P_i = a - bQ_i - btQ_j, \qquad a>0,\ b>0,$$

where P_i is the (net) price, Q_i is the quantity sold by firm i, a is the demand intercept, b is the slope of the demand curve, and $t (0<t\leq 1)$ represents the degree of substitutability between products i and j.[10] Let $b=1$ without loss of generality, by normalizing Q_i and Q_j appropriately. With constant and identical marginal costs of production for both firms, P_i is defined as net of the marginal costs.

To maximize profit of $\Pi_i = P_iQ_i = (a-Q_i-tQ_j)\cdot Q_i$, firm i sets $2Q_i = a - tQ_j$ to satisfy the first-order condition.[11] This derives firm i's reaction function as $Q_i = a/2 - t/2Q_j$, which is negatively related to Q_j. The unique equilibrium output is then $Q_i = a/(2+t)$, and the equilibrium price is $P_i = a/(2+t)$. Substituting the first-order condition into the profit function yields $\Pi_i = Q_i^2$: firm i's profit is its equilibrium output squared. This further suggests that expected profits are convex in the expected equilibrium quantity under uncertainty. When faced with a choice on disclosure (and subsequent output choice), firms will choose the option that will maximize variance in equilibrium outputs.

A similar analysis can be carried out for Bertrand competition. Rewrite the demand function as $Q_i = \alpha - \beta P_i + \beta t P_j$, where $\alpha = a/(1+t)$ and $\beta = 1/(1-t^2)$. The first-order condition yields the reaction function as $P_i = \alpha/2\beta + (t/2)P_j$, which is positively related to P_j. This unique price and output of firm i are $P_i = \alpha/\beta(2-t) = a(1-t)/(2-t)$ and $Q_i = \alpha/(2-t) = a/(1+t)(2-t)$. Substituting the first-order condition in the profit equation yields $\Pi_i = \beta P_i^2$. Again, this suggests that expected profits will be convex in the equilibrium price under uncertainty. The two features that are useful in understanding the ensuing analysis under uncertainty are that (1) the reaction functions are sloped negatively (positively) and (2) the equilibrium profits are convex in the equilibrium quantity (*net* price) in Cournot (Bertrand) competition.

---

[8] In the subsequent analysis, it turns out that firms will choose either no noise or infinite noise. Thus, henceforth, the word "disclosure" will mean disclosure of private information with no noise and "no disclosure" will mean disclosure with infinite noise.

[9] Since I am interested in symmetric equilibrium, I will present equations for only one firm to save space. When clear, I omit references to the indices, e.g., $i\neq j=1,2$.

[10] Substituting equilibrium quantities and prices below, one can compute cross-price elasticity of demand at equilibrium as $t/(1-t^2)$ for Cournot, and t for Bertrand, competition.

[11] The analysis for firm j is omitted, since it is completely symmetrical.

Uncertainty

In introducing uncertainty about demand or cost, I rely heavily on the techniques of Gal-Or (1985, 1986).[12] Industry demand is stochastic as:

$$P_i = a + \Delta a - Q_i - tQ_j, \qquad a>0,\ 0<t\leq 1, \qquad i\neq j=1,2,$$

where Δa is a stochastic disturbance in the demand intercept. Again, the value of a is interpreted as *net* of the constant (and known portion of) marginal costs. In addition, marginal costs can have a stochastic portion, Δc_i. Although the source of uncertainty is twofold, I analyze each separately below. Further assumptions are made as follows. Demand uncertainty Δa_i is normally distributed with zero mean and variance of $\sigma > 0$. Each firm receives an imperfect signal, $x_i = \Delta a_i + e_i$, with $e_i \sim N(0, m)$ and $m \geq 0$. The aggregate signal is the average $\Delta a = (\Delta a_1 + \Delta a_2)/2$ with $\Delta a \sim N(0, \sigma/2)$.

Each firm decides how to disclose its signal by choosing a noise level in the message. The firm's message is denoted as $\hat{x}_i = x_i + f_i$, with $f_i \sim N(0, s_i)$. The variance terms in the signal and the message may be interpreted as their "quality" or "precision." A similar information structure is assumed for costs. Marginal costs are distributed normally, $\Delta c_i \sim N(0, \sigma_c)$. After receiving a signal, $z_i = \Delta c_i + \epsilon_i$ with $\epsilon_i \sim N(0, n)$, each firm reports $\hat{z}_i = z_i + g_i$ with $g_i \sim N(0, v_i)$. Throughout the analysis, signals (x_i and z_i) and costs (Δc_i) are assumed to be independent and identically distributed. Note that stochastic costs are firm-specific (private values), whereas stochastic demand is common to both firms (common values); all the results on demand and cost hinge on this difference.

The assumptions of linear inverse demand and normal distributions allow computation of expectations to derive the unique equilibria.[13] With this basic structure, four cases with different market and information configurations are analyzed: Cournot versus Bertrand for market competition, and demand versus cost for private information.[14]

Demand Uncertainty

In a Cournot duopoly game, firms are choosing quantities, and the precise value of the demand intercept is uncertain ($\sigma > 0$). Each firm is concerned with disclosure (how much noise to add) and quantity choice. Given the aggregate information structure, σ, m, and the firm-specific information x_i, the strategy choice is then a pair (s_i, Q_i). The game is played sequentially: firms decide first on disclosure and then on quantity. In the two scenarios, firms either commit to a disclosure strategy before receiving private signals (*ex ante* scenario), or decide on disclosure after receiving their respective signals (*ex post* scenario).

---

[12] The basic model is a synthesis of models developed in the literature on information sharing, particularly those by Gal-Or. Gal-Or (1985) investigates Cournot competition with demand uncertainty in a homogeneous product market, whereas Gal-Or (1986) investigates both Cournot and Bertrand competition with cost uncertainty.

[13] In particular, the assumption of normal distribution is convenient for obtaining linear posterior values. Other density functions, such as the gamma or beta functions, also yield linear posteriors.

[14] With the assumption of constant marginal costs, it is possible to have equilibria in which one firm is choosing quantity and the other price, but the present analysis is restricted to more familiar Cournot and Bertrand competition. See Klemperer and Meyer (1986) for an analysis of how the strategic variables chosen by firms are affected by the shape of marginal costs and the nature of demand uncertainty.

Since each firm makes decisions sequentially, it is logical to analyze the second stage first. This guarantees a subgame perfect equilibrium such that firms will have no incentive to deviate once they reach the second stage. At that stage, each picks an optimal quantity level, given s_1 and s_2 (committed to and made public in the first stage). The unique Bayesian-Nash equilibrium output is a linear function of private signal x_i and disclosed messages ($\hat{x}_i$ and $\hat{x}_j$) as:[15]

$$Q_i=\frac{1}{(2+t)}\left[a+E(\Delta a_j|\hat{x}_j)-\frac{t}{4}E(\Delta a_i|\hat{x}_i)+\frac{(2+t)}{4}E(\Delta a_i|x_i)\right],\qquad i\neq j=1,2, \quad (1)$$

where $E(\bullet\,|\,\bullet)$ is the expectation operator. By substituting posterior distributions, the Bayesian-Nash equilibrium quantity is formally stated in the following lemma.

Lemma 1: For given (precommitted) s_1 and s_2, the unique Bayesian-Nash equilibrium quantity for firm *i* in a Cournot duopoly game with demand uncertainty is:

$$Q_i=\frac{1}{(2+t)}\left[a+\frac{\sigma}{2(m+\sigma+s_j)}\hat{x}_j-\frac{t\sigma}{4(m+\sigma+s_i)}\hat{x}_i\right.$$
$$\left.+\frac{\sigma(2+t)}{4(m+\sigma)}x_i\right],\qquad i\neq j=1,2. \quad (2)$$

A close inspection of equation (2) reveals insights that will be useful for later analysis. Obviously, without any private signal, the quantity chosen will be identical to that under certainty (quantity of $a/(2+t)$ since $E(\Delta a)=0$). If the message from firm *j* is infinitely noisy, firm *i* will adjust its quantity by using only its private information. Sending a message affects *i*'s own equilibrium quantity since *j* would respond to the message. This process is depicted in figure 1 (top panel). Without any signal, both firms expect a market equilibrium at *E*. Suppose now that firm 1 has received a favorable signal ($x_1>0$), while firm 2 has not received any signals at all. Firm 1 will shift its reaction curve to the right. If firm 1 committed (in the first stage) not to send any message (or send a message with infinite noise), firm 2 will plan to produce Q_2^{ND}. The expected equilibrium, then, is the point depicted by *ND*. Alternatively, if firm 1 were to send a message (say, with no noise), firm 2 would update the expected value of Δa upward and shift its reaction curve upward.[16] The ensuing equilibrium is depicted by point *D*. As equation (2) shows, *D* is expected to be to the left of *ND* (*i.e.*, $Q_1^D<Q_1^{ND}$).

An important observation is that the equilibrium quantity of a firm, *ceteris paribus*,

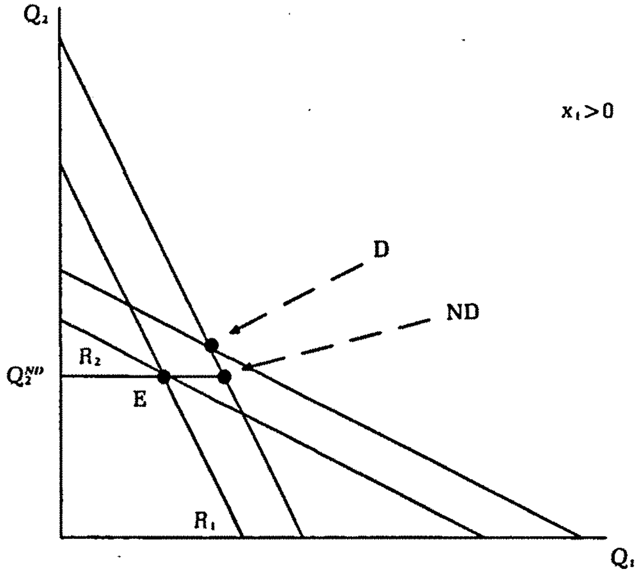
---

[15] With the assumption of linear demand and normally distributed signals, it is possible to have nonpositive quantities or prices. Typically, the literature ignores this possibility by assuming relatively small variance so that such an event becomes unlikely. See Novshek and Sonnenschein (1982) and Vives (1984).

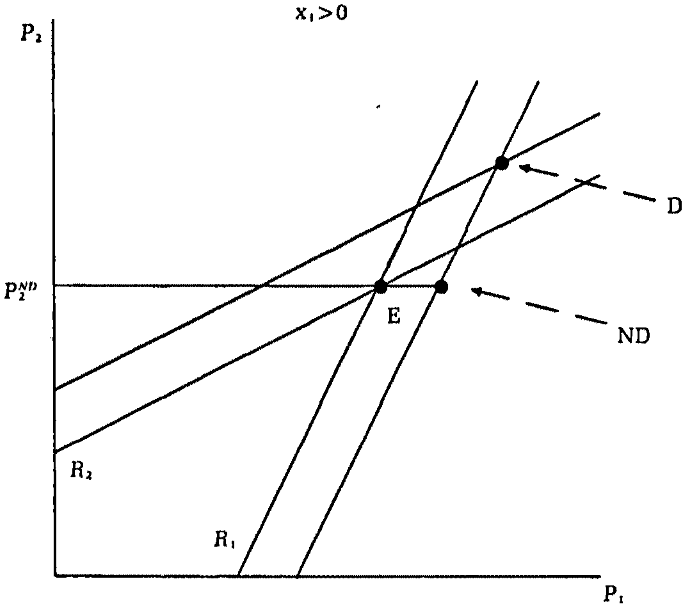
[16] How much reaction curves shift depends on the accuracy of signals and messages in relation to the underlying variability of demand. The more accurate signals become ($m\rightarrow 0$), the larger the effect. In fact, if signals are perfect ($m=0$) and firms disclose without noise ($s_1=s_2=0$), both Q_1 and Q_2 will adjust fully (by $E(\Delta a|x_1,x_2)=(x_1+x_2)/2$).

Figure 1
Ex Ante Equilibrium with Demand Uncertainty

Cournot



Bertrand



is positively correlated to its own signal but negatively correlated to its message.[17] This is because a disclosure induces the rival firm to respond in the same direction, which reduces the first firm's equilibrium quantity. The two products are strategic substitutes in Cournot competition.

Since the response to a private signal is larger than that to the message sent, the overall effect is a positive correlation between two firms' outputs. Without any disclosure, however, firm outputs are uncorrelated.

Positive correlation implies that disclosure will increase j's output at the expense of i's output, when i's signal is favorable. Absent disclosure, firm i would be able to increase its output even more. Recall that expected profits under uncertainty are convex in the equilibrium quantity. Hence firm i can increase its expected profit by remaining silent about its private signal. This implies that, in choosing a disclosure policy *ex ante*, no disclosure (or infinitely noisy disclosure) will be the Bayesian-Nash equilibrium strategy. Formally:

> *Proposition 1*: The unique Bayesian-Nash equilibrium of the two-stage Cournot game with demand uncertainty is a pair, $s_i = \infty$ and Q_i, which satisfies equation (2).

This result is easy to interpret in figure 1. The fact that Q_1^{ND} is larger then Q_1^{D} when $x_1 > 0$ implies that the variability of equilibrium quantity is higher without disclosure. Since expected profits are convex in the equilibrium quantity, a policy that provides higher variability is optimal.[18] It should be noted that no disclosure forms a dominant strategy *ex ante*. This suggests that firms might be reluctant to support any mandatory disclosure policy involving information on uncertain demand.

When $s_1 = s_2 = \infty$, expected profit is:

$$E\Pi_i = \frac{1}{(2+t)^2}\left[a^2 + \frac{\sigma^2(2+t)^2}{16(m+\sigma)}\right]. \tag{3}$$

This shows that expected profits are larger than those under certainty ($\sigma = m = 0$), in which case, $\Pi_i = a^2/(2+t)^2$. This obtains because, when demand is stochastic, firms are able, by using available information, to exploit fluctuations in demand; the increased profits when demand is favorable more than compensate for the reduction in profits when demand is unfavorable. As signals get better ($m \rightarrow 0$), firms' expected profits, of course, increase.

When products are not good substitutes, it is interesting to note that expected profits are higher if both firms always disclose, say, with zero noise. In such a case,

$$E\Pi_i = \frac{1}{(2+t)^2}\left[a^2 + \frac{\sigma^2}{2(m+\sigma)}\right],$$

which is strictly larger than the expected profit under no disclosure as in equation (3) as long as, $t < 2\sqrt{2} - 2$, or about 0.828. In other words, if firms can negotiate a binding agreement for disclosure they are better off, as in the "prisoners' dilemma." This, however, is not a viable agreement in a noncooperative setting. When the mechansism of information transmission is a voluntary organization, such as a trade association, it is difficult to imagine that the association can enforce such an agreement. If, however, the mechanism is either the FASB or the SEC, the enforcement power is substantially greater. Since these agencies can penalize firms that do not comply with the regulation, firms facing demand uncertainty in a higher differentiated Cournot industry might have an incentive to lobby for mandating disclosure of demand-related information, even though they would choose not to share the same information through trade associations.

---

[17] This is the case regardless of the signal received or the message sent by the other firm.

[18] This suggests that, when firms are risk-averse, the higher variability is not unambiguously preferred.

In a Bertrand game, firms are choosing prices rather than quantities. All the parameters are identical to the Cournot game. (Again assume that costs are certain.) The following proposition summarizes the equilibrium first-stage choice of disclosure and the second-stage choice of price.

Proposition 2: The unique Bayesian-Nash equilibrium of the two-stage Bertrand game with demand uncertainty is a pair, $s_i=0$ and P_i which satisfies:

$$P_i=\frac{(1-t)}{(2-t)}\left[a+\frac{\sigma}{(m+\sigma+s_j)}x_j+\frac{t\sigma}{2(m+\sigma+s_i)}x_i+\frac{(2-t)\sigma}{4(m+\sigma)}x_i\right],\qquad i\neq j=1,2. \tag{4}$$

It is immediately clear that the expected profit decreases with the value of s_i. Again, the optimal strategy forms a dominant strategy.

Notice that a firm's responses to its private signal and to its own message are positively correlated. This implies that disclosure increases the variability of the equilibrium price. Hence, noiseless disclosure forms a dominant strategy. Positive correlation follows from the fact that the two firms are strategic complements. Thus, while prospects of higher demand increase prices, the disclosure of favorable information *raises* the prices of *both* firms even more and results in higher expected profits. This is depicted in the bottom panel of figure 1. Again, assume that firm 1 received a favorable signal. Without disclosure, firm 2 would expect to price at P_2^{ND}; with a favorable message, firm 2 shifts its reaction curve upward. The resulting expected equilibrium prices are depicted at *D*. Since expected profits are convex in equilibrium (net) prices, noiseless disclosure increases expected profits by increasing the variability of equilibrium prices.

In contrast to Cournot, Bertrand competition is more "intense" as it drives prices down. This can be easily seen from the fact that, if products were perfect substitutes ($t=1$), then expected (net) prices and profits would be identically zero. The only reason for generating strictly positive expected profits is product differentiation. Cournot competition does not drive prices down to the same extent. Although the Cournot prices are lower than the monopoly price (because two firms might have colluded), they are still higher than the Bertrand prices. Given the cut-throat nature of Bertrand competition, mutually providing more accurate information benefits firms only when demand is uncertain. Otherwise, firms will miss the opportunities to set prices higher when demand is high. Firms are better able to coordinate pricing behavior and reduce the damaging effect of severe competition by sharing information.

Cost Uncertainty

Now assume that demand is certain ($\sigma = m = 0$) but marginal costs are random, firm-specific, and independently distributed. Although an exactly analogous analysis applies to the second-stage quantity and price choices, the first-stage choice is now reversed. *Ex ante*, firms would commit to noiseless disclosure under Cournot, but infinitely noisy disclosure under Bertrand.

Proposition 3: The unique Bayesian-Nash equilibrium strategy of the two-stage Cournot game with cost uncertainty is a pair, $v_i = 0$ and Q_i, which satisfies,

$$Q_i = \frac{1}{(2-t)}\left[a + \frac{t\sigma_c}{(2-t)(n+\sigma_c+v_j)}\hat{z}_j - \frac{t^2\sigma_c}{2(2-t)(n+\sigma_c+v_i)}\hat{z}_i - \frac{(2+t)\sigma_c}{2(n+\sigma_c)}z_i\right]. \tag{5}$$

Similarly when firms are playing Bertrand, we have:

Propositon 4: The unique Bayesian-Nash equilibrium strategy of the two-stage Bertrand game with cost uncertainty is a pair, $v_i = \infty$ and P_i, which satisfies:

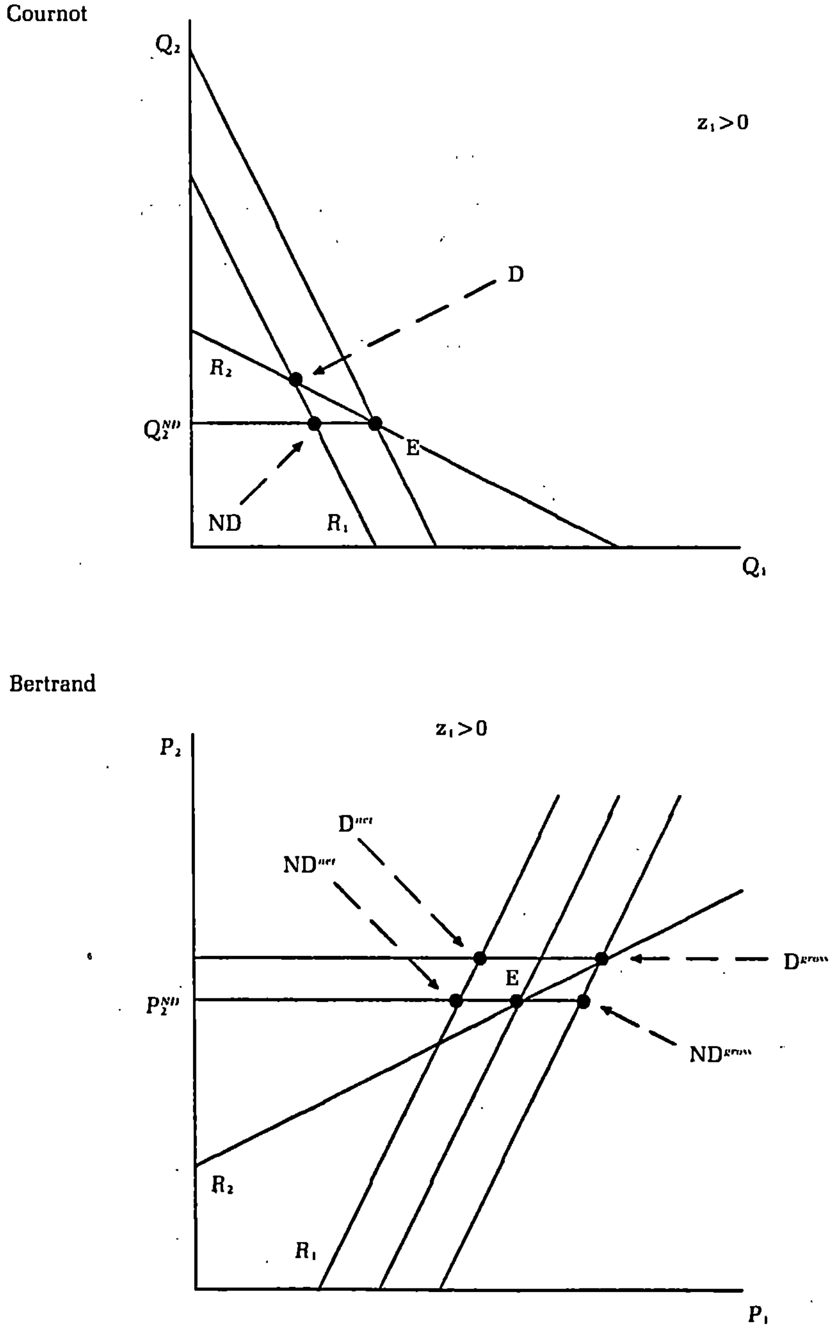
$$P_i = \frac{(1-t)}{(2-t)}\left[a + \frac{t\sigma_c}{(2+t)(1-t)(n+\sigma_c+v_j)}\hat{z}_j + \frac{t^2\sigma_c}{2(2+t)(1-t)(n+\sigma_c+v_i)}\hat{z}_i - \frac{(2-t)\sigma_c}{2(1-t)(n+\sigma_c)}z_i\right], \tag{6}$$

where P_i is net of expected marginal costs.

Again, opposite disclosure incentives obtain when costs are certain. To understand why uncertainty about either demand or cost makes a difference in resulting equilibria, recall that demand is a common value, whereas costs are private values. In fact, a higher demand intercept can be viewed as equivalent to a lower (constant) marginal cost. What is important, therefore, is not demand or cost uncertainty *per se*, but rather how uncertain parameters are incorporated in decision making. The specific assumption about demand is additivity, which in effect forces firms to become interdependent.

The intuition behind Propositions 3 and 4 is straightforward. In both cases, absent disclosure (or with infinitely noisy disclosure), the choice variables are uncorrelated, as before. When signals are transmitted, the equilibrium quantity is further reduced (increased) if a signal is unfavorable (favorable), which causes larger (smaller) variance in equilibrium quantity. Positive correlation here is attributed to the fact that the other firm responds better to the cost condition when information is less noisy (by moving along the reaction curve), but it cannot respond without disclosure. Thus, when firm 1's cost is high, its output falls more because firm 2 increases its output, and so on. This is depicted in the first panel of figure 2.

Figure 2
***Ex Ante* Equilibrium with Cost Uncertainty**



Cournot

Bertrand

With Bertrand competition, however, the coefficients on the private signal received and the firm's own message sent are negatively correlated. Disclosure will dampen the disclosing firm's response to a signal and reduces the variability of equilibrium net price. For example, the receipt of an unfavorable signal $(z_1 > 0)$ will shift firm 1's reaction curve to the right (starting with the curve in the middle). Without disclosure, firm 2 will price at P_2^{ND} to result in an equilibrium at ND^{gross}. With disclosure, however, an equilibrium will obtain at D^{gross}. Minus $E(\Delta c_1/z_1)$, the net prices firm 1 receives, however, would be ND^{net} and D^{net}. It is clear that no disclosure yields higher variability in net prices. Thus, when firms have the option to commit to a disclosure policy *ex ante*, they would rather not pool the information. This is true even if firms were able to come to a binding agreement for disclosure. It turns out that the effect of disclosure on expected profit is negative when $v_j = v_i < \infty$. Even if a mandatory disclosure rule guaranteed no defection, firms would not want information sharing, quite contrary to the situation in the Cournot/demand case. Again, firms will have different preferences *ex post*, after the realization of their private signals.

The *ex ante* preferences of firms are summarized below. Firms will choose infinite noise (*ND*, no disclosure) as a dominant strategy in the cases of Cournot/demand and Bertrand/cost. Similarly they will choose zero noise (*D*, full disclosure) in Cournot/cost and Bertrand/demand cases. Recall, however, that firms are better off if they can bind themselves into disclosure in the case of Cournot/demand (when products are highly differentiated), but not in the case of Bertrand/cost.

	Competition	
	Cournot	*Bertrand*
Demand	*ND*	*D*
Cost	*D*	*ND*

Ex ante preferences are chosen by considering the effect of disclosure on expected profits. Since firms are identical *ex ante* and the distribution of signals is common knowledge, the financial market will be able, on average, to value firms correctly. The potential tension between the financial-market valuation and product-market competition does not exist in the *ex ante* setting. Such a tension, however, might arise in the ex *post* setting once a firm receives a signal and considers the effect of its disclosure on the two markets.

III. *Ex Post* Setting

In the absence of a precommitted disclosure policy, whether a firm finds it in its best interest to disclose and communicate private signals partially depends on how lack of disclosure is interpreted by the competitor. In the *ex ante* model, a firm adheres to a particular disclosure policy regardless of the realized value of the signal. Therefore, even if there is no disclosure, the very act of nondisclosure does not provide any information as to what type of private signal a firm has received. In the *ex post* scenario, however, upon observing nondisclosure by a firm, the other firm rationally asks why a firm would not disclose when it may do so. One possibility is that the firm has not received any signal. Another possibility is that the firm has received a signal of the type that is better undisclosed.

The assumption in this section is that firms do not always receive a signal, but the market has (common) prior expectations about the likelihood of firms' receiving signals. In such a setting, it is useful to analyze the ex *post* model as a convex combination of two polar cases: one in which firms expect to receive private signals for sure, the other in which firms do not expect to receive any signal. In the first case, nondisclosure is interpreted as private information of certain types; in the second, nondisclosure is attributed to lack of information. In an intermediate case, a firm may suspect that the other firm received a signal, but cannot be sure. Nondisclosure is then attributed partially to lack of private information and to the receipt of the type of signal that is better undisclosed.

Again, the setting is duopolists competing in a two-stage game. In the first stage, each firm receives a signal about demand or cost with probability, $\gamma_i = \gamma_j = \in [0,1]$.[19] This prior probability is assumed to be common knowledge. If a firm receives a signal, it then decides whether to disclose it to the other firm. In the second stage, firms decide output (price) levels, taking into account the private signals they may have received as well as publicly disclosed messages, or the absence thereof, in the first stage. The two periods can be viewed as any two consecutive periods in which information transmission and production take place sequentially in an infinite-period game. If demand or cost is stochastic, and independently and identically distributed in each time period, there is nothing firms can learn over time. Thus, it is sufficient to investigate only two periods.

In the *ex post* setting, it is straightforward to show that the unique equilbrium is full disclosure in Cournot/cost and Bertrand/demand combinations, when the precommitment policy is full disclosure. Firms with more favorable signals are better off with disclosure, while firms with more unfavorable signals want to hide their information. Because a nondisclosing firm will be considered the average of nondisclosing firms, those nondisclosing firms that are better than the average, however, will want to disclose their private information rather than be judged worse than average. Hence, nondisclosure unravels. This is precisely the situation of full revelation examined in Grossman (1981) and Milgrom (1981). Only in a limiting case in which the market believes that there is absolutely no private information, will nondisclosing firms be considered as average. Thus, in these two configurations of competition and information, full disclosure is the unique equilibrium in both *ex ante* and *ex post* settings.

Interesting cases are Cournot/demand and Bertrand/cost configurations. Without precommitment to the nondisclosure, firms find it difficult to withhold information. Signals that firms want to disclose *ex post* might bring pressure for further disclosure. Yet, nondisclosure need not unravel completely in the ex *post* setting because firms with nonextreme signals would disclose, but firms with extreme (both favorable and unfavorable) signals would hide behind nondisclosure. Then a nondisclosing firm would be attributed with the average of signals that firms want withheld. A firm with a signal equal to the average would be indifferent between disclosure and nondisclosure, which implies a sustainable equilibrium.

To simplify the analysis, assume that (1) signals are perfect ($m = n = 0$); (2) the disclosure strategy of firms is dichotomous, with two pure strategies, $d \in (D, ND)$; and (3) any disclosure (D) is noiseless, accurate, and credible. Since disclosure with infinite noise is equivalent to absence of disclosure, I refer to this as nondisclosure (ND).

---

[19] The more general case where $\gamma_i \neq \gamma_j$ is considered by Sankar (1992).

Cournot with Demand Uncertainty

A first conjecture is that firms will disclose only some values of $x_i \in X_i^D(\gamma)$ and will not disclose $x_i \in X_i^{ND}(\gamma)$, $0<\gamma\leq 1$.[20] Similarly, firm *j* will disclose $x_j \in X_j^D(\gamma)$ and not disclose $x_i \in X_j^{ND}(\gamma)$. Let $\phi_i(\gamma)=\text{Prob}[x_i \in X_i^{ND}(\gamma)]$ and $\phi_j(\gamma)=\text{Prob}[x_j \in X_j^{ND}(\gamma)]$.

Firm *i* will produce the following output in the second stage (See Lemma 1).[21]

$$Q_i^d=\frac{1}{(2+t)}\left[a+\frac{1}{2}\bar{x}_j-\frac{t}{4}(\bar{x}_i-x_i)+\frac{1}{2}x_i\right]. \tag{7}$$

Expected profits for firm *i* are:

$$E\Pi_i^d=(Q_i^d)^2$$

$$=\frac{1}{(2+t)^2}\left[a+\frac{1}{2}\bar{x}_j-\frac{t}{4}(\bar{x}_i-x_i)+\frac{1}{2}x_i\right]^2, \tag{8}$$

where $Q_i^d, d\in(D,ND)$ is the output chosen after disclosure or nondisclosure, given the private signal and the (possible) messages to and from firm *j*, $\bar{x}_i$, and $\bar{x}_j$. For most values of signals and messages, Q_i^d are nonnegative; however, since the distribution of x_i and x_j are unbounded, it is possible to have a negative quantity giving rise to a positive profit. This is possible, for example, if a firm receives a subsidy from the government for nonproduction.[22] In equation (8), $\bar{x}_i=x_i$ when firm *i* discloses its signal. If there is no disclosure, $\bar{x}_i$ is the expected value of $x_i \in X_i^{ND}(\gamma)$, and is also denoted by $E[x_i|x_i \in X_i^{ND}(\gamma)], \gamma \in [0,1]$. To see how the expected value changes with γ, consider first the case of $\gamma=1$.[23] (The more general case of $0<\gamma<1$ will follow.)

Expecting a receipt of a signal, x_i, firm *i* formulates an equilibrium disclosure policy, $X_i(\gamma)\in[X_i^D(\gamma), X_i^{ND}(\gamma)]$, by selecting the disclosure region $X_i^D(\gamma)$ and the nondisclosure region $X_i^{ND}(\gamma)$ to maximize expected profit, given by equation (8). Let $\gamma=1$. Any Bayesian-Nash equilibrium must satisfy the condition that, given firm *j*'s disclosure strategy $X_j(1)$, firm *i* chooses $X_i(1)$ such that:

$$E\Pi_i^D[x_i|X_j(1)]\geq E\Pi_i^{ND}[x_i|X_j(1)] \qquad \forall x_i \in X_i^D\,(1),$$

$$E\Pi_i^D[x_i|X_j(1)]\leq E\Pi_i^{ND}[x_i|X_j(1)] \qquad \forall x_i \in X_i^{ND}(1), \tag{9}$$

where $E\Pi_i^D(\bullet\,|\,\bullet)$ is firm *i*'s expected profit from disclosure and $E\Pi_i^{ND}(\bullet\,|\,\bullet)$ from nondisclosure. Given the disclosure policy of the other firm *j*,

---

[20] This analysis is based on the model developed by Clinch and Verrecchia (1991), in which Cournot/demand configuration is analyzed with $\gamma=1$.

[21] The upper bars in the following equations indicate both disclosed and expected values of private information.

[22] Another possible interpretation of negative quantity is that the producing firm becomes a net consumer of the product.

[23] It should be noted that in this case the existence of nondisclosure regions depends crucially on a positive probability of negative production that is beneficial to the firm. This will be discussed further subsequently.

$$E\Pi_i^D(\bullet\,|\,\bullet)=\phi_j(1)\left(a+\frac{1}{2}x_i+\frac{1}{2}\bar{x}_j\right)^2$$

$$+[1-\phi_j(1)]E\left[\left(a+\frac{1}{2}x_i+\frac{1}{2}x_j\right)^2|x_j\in X_j^D(1)\right] \text{ and}$$

$$E\Pi_i^{ND}(\bullet\,|\,\bullet)=\phi_j(1)\left(a+\frac{1}{2}x_i+\frac{1}{2}\bar{x}_j-\frac{t}{4}(\bar{x}_i-x_i)\right)^2$$

$$+[1-\phi_j(1)]E\left[\left(a+\frac{1}{2}x_i+\frac{1}{2}x_j-\frac{t}{4}(\bar{x}_i-x_i)\right)^2|x_j\in X_j^D(1)\right],$$

where $\phi_j(1)$ is the probability that firm j does not disclose its signal when $\gamma=1$. Recall that (truthful) disclosure implies $\bar{x}_i=x_i$, but without disclosure $\bar{x}_i=E[x_i|x\in X_i^{ND}(1)]$. This further implies:

$$E\Pi_i^{ND}(\bullet\,|\,\bullet)-E\Pi_i^D(\bullet\,|\,\bullet)=\phi_j(1)\left[\frac{t^2}{16}(\bar{x}_i-x_i)^2-\frac{t}{2}(a+\frac{1}{2}x_i+\frac{1}{2}\bar{x}_j)(\bar{x}_i-x_i)\right]$$

$$+[1-\phi_j(1)]E\left[\frac{t^2}{16}(\bar{x}_i-x_i)^2-\frac{t}{2}(a+\frac{1}{2}x_i+\frac{1}{2}x_j)\right.$$

$$\left.(\bar{x}_i-x_i)|x_j\in X_j^D(1)\right]$$

$$=(\bar{x}_i-x_i)\left(\frac{t^2}{16}(\bar{x}_i-x_i)-\frac{t}{2}(a+\frac{1}{2}x_i)-\frac{t}{2}\left\{\phi_j(1)\bullet\frac{1}{2}\bar{x}_j\right.\right.$$

$$\left.\left.+[1-\phi_j(1)]E\left[\frac{1}{2}x_j|x_j\in X_j^D(1)\right]\right\}\right)$$

$$=\frac{t}{2}(\bar{x}_i-x_i)\left(\frac{t}{8}\bar{x}_i-a-\frac{t+4}{8}x_i\right),$$

since,

$$\phi_j(\gamma)\bar{x}_j+[1-\phi_j(\gamma)]E[x_j|x_j\in X_j^D(\gamma)]=0 \qquad \forall\gamma.$$

The following proposition from Clinch and Verrecchia (1991) summarizes the equilibrium.

Proposition 5: In Cournot competition with demand uncertainty, the Bayesian-Nash equilibrium strategy of firm i in the *ex post* setting in which firms are expected to possess private information is characterized by:

$$X_i^D(1)=\left(\frac{t\bar{x}_i-8a}{t+4},\bar{x}_i\right),$$

where $\bar{x}=E[x_i|x_i\in X_i^{ND}(1)]$ and $0<\bar{x}_i<8a/t$.[24]

A firm will disclose a signal from the interval, but will not disclose a signal from below or above the interval. In other words, relatively extreme values will not be transmitted. A firm chooses whether to disclose by comparing expected profits with and without disclosure. Without disclosure, a firm will be regarded as an average firm of the nondisclosing group. Notice that a firm with $x_i = \bar{x}_i$ will be indifferent between disclosure and nondisclosure. If the nondisclosure region were a single interval, the expected value could not be at either end of the region. For any nondisclosure region to exist then, it has to be accompanied by at least one more nondisclosure region on a real line.

Note that this particular disclosure region is based on the assumption that firms sometimes produce negative quantities that result in positive profits. Since this assumption is somewhat difficult to interpret, it might be more acceptable to restrict quantities to be nonnegative. It has been shown by Sankar (1992) that a similar nondisclosure region exists at the lower tail, if we assume that firms choose nondisclosure over disclosure when output is zero.[25] Thus, the qualitative results are the same even when negative output is precluded.[26]

Clinch and Verrecchia (1991) show that such an equilibrium exists and that the disclosure interval decreases with the value of t. Since their model does not assume any density function for the signal, it is not possible to estimate the size of this interval. Assuming normally distributed x_i, I compute the interval for arbitrary chosen values of a and t. For example, let $x_i \sim N(0,1)$ and $a=5$. The interval, which depends on t, is presented below. Note that the area is "virtually" 100 percent of all values of t.

t	*Lower Threshold*	*Upper Threshold*	*Interval*	*Area (%)*
0.7	-7.78	4.90119	12.68	100
0.9	-7.26	4.90116	12.16	100
1.0	-7.02	4.90109	11.92	100

The figures show the lower threshold, the upper threshold, the interval between the two thresholds, and the cumulative area of the disclosure region. Notice that the upper threshold is the expected value of the regions outside the interval. Virtually all values of x_i will be disclosed. For a smaller value of a, the interval does get marginally smaller.[27]

Given this ex post equilibrium disclosure strategy, would firms prefer to precommit to nondisclosure *ex ante*? It is straightforward to show that firms would precommit to

---

[24] The necessary condition follows from the fact that $\bar{x}$ is the expected value of the nondisclosure regions. If $\bar{x}<0, X_i^p(\cdot)$ will consist of only negative values. Then $\bar{x}$ will be positive, a contradiction. Similarly, if $\bar{x}>8a/t$, $X_i^p(\cdot)$ consists of only positive values, $\bar{x}$ will be negative, also a contradiction.

[25] Clearly, it is essential to have at least two nondisclosure regions to sustain nondisclosure in equilibrium. In the Cournot/cost and Bertrand/demand cases, nondisclosure is preferred by firms with mid-range signals (when negative output is allowed) or below some point (when negative output is precluded), which results in at most one potential nondisclosure region. Thus, this nondisclosure region unravels in a rational-expectation equilibrium.

[26] As the lower threshold shifts down, the upper threshold shifts up, which causes a larger disclosure region. The case of nonnegative output is analyzed in detail by Sankar (1992). She investigates a generalized model of disclosure in Cournot oligopoly when firms might receive signals on both firm-specific and industry-wide cost information.

[27] For example, with $a=2$ and $t=0.9$, the area becomes 97.3 percent.

nondisclosure when, $t>2\sqrt{2}-2$, as in the case in the *ex ante* analysis.[28] Thus, nondisclosure is not subgame perfect without precommitment.

The analysis so far shows that it is virtually impossible to hide private information when $\gamma=1$. If $\gamma<0$, however, it becomes possible to withhold information, since nondisclosure can be attributed to lack of information. To see how equilibrium behavior is affected by γ, consider the more general case of $0<\gamma<1$, in which firms are not sure whether the competitor has received a signal. Thus, if firm *j* does not disclose, firm i has to make a conjecture about what signal firm *j* has received, if any. If firm *i* conjectures that firm *j* has not received a signal (as if $\gamma=0$), firm *i* will attribute firm *j* with the prior (expected) value, $E[x_i|x_i \in X_i^{ND}(0)]=0$. Alternatively, if firm *i* conjectures that firm *j* received a signal (as if $\gamma=1$) but has chosen not to disclose, then firm i will conclude that the signal was of the type that is better withheld. In fact, firm *j* has received a signal with probability γ. Thus, when $0<\gamma<1$, nondisclosure will be attributed with a posterior (expected) value, which is the weighted average of the expected values of the two polar cases. Thus,

$$E(x_i|ND,\gamma)=\frac{1-\gamma}{1-\gamma+\gamma\phi_i(\gamma)}E[x_i|x_i \in X_i^{ND}(0)]+\frac{\gamma\phi_i(\gamma)}{1-\gamma+\gamma\phi_i(\gamma)}E(x_i|x_i \in X_i^{ND}(1)]$$

$$=\frac{\gamma\phi_i(\gamma)}{1-\gamma+\gamma\phi_i(\gamma)}E[x_i|x_i \in X_i^{ND}(1)]>0, \qquad 0<\gamma\leq 1.$$

What happens to the incentive to disclose as $\gamma\to 0$? Clearly, as $\gamma\to 0, E(x_i|ND,\gamma)\to 0$. This implies that the upper nondisclosure region is enlarged to include virtually all positive values. The lower nondisclosure region, however, moves downward. A straightforward calculation shows that as $\gamma\to 0$, the disclosure region approaches:[29]

$$\lim_{\gamma\to 0} X_i^D(\gamma)=\left[\frac{-8a}{t+4},0\right].$$

Thus, the disclosure region is almost the entire nonpositive region in the case of Cournot with demand uncertainty. Similar results hold for the other competition/information configurations: ex post, the disclosure incentives diverge according to whether the signal is above or below the prior expected value. For firms with favorable signals, the optimal disclosure policy is identical to the *ex ante* policy. This is summarized in the following proposition.

Proposition 6: When γ is arbitrarily close to zero, a firm that has received a favorable signal would follow the same disclosure policy as in the *ex ante* setting with precommitment, whereas a firm that has received an unfavorable (but not extremely unfavorable) signal would follow the opposite disclosure policy.

The intuition is straightforward. It was shown in the previous section that firms would choose a policy that yields higher variability in the expected equilibrium outputs (or net prices) so as to maximize expected profits. Once a signal is realized, however, firms are better off with a policy that yields a higher expected output (or net price). For a firm with a favorable signal ($x_i > 0$ or $z_i < 0$), it is the same policy as in the *ex ante* choice; for a firm with an unfavorable signal (except for extremely unfavorable signals), it is the opposite. If firms receive information that is not routinely expected by the market, the present analysis shows that "bad news" tends to be released in the Cournot/demand and Bertrand/cost cases, while "good news" tends to be released in the Cournot/cost and Bertrand/demand cases.

---

[28] Recall that *ex ante* nondisclosure is a dominant Nash Strategy, but when, $t<2\sqrt{2}-2$, firms are better off if they have a binding agreement to disclosure.

[29] A numerical calculation shows that the disclosure region, $1-\phi(\gamma)$, changes with the value of γ as follows: $1-\phi(0.8)=0.907$; $1-\phi(0.5)=0.797$; $1-\phi(0.3)=0.708$; $1-\phi(0.1)=0.585$. These values are insensitive to the value of t. It is not feasible to characterize an equilibrium disclosure region analytically.

This analysis offers an interesting insight for the precommitment model, by suggesting conditions under which a firm has incentive to undermine the precommitted policy once an unfavorable signal is realized. Even though firms have precommitted to nondisclosure, those firms that have received unfavorable signals could make themselves better off if they could defect from their precommitted policy. If the policy is being implemented by a trade asssociation, firms with unfavorable cost signals in Cournot and unfavorable demand signals in Bertrand would, if possible, terminate their membership to prevent signal dissemination.[30] Alternatively, if disclosure is mandatory and governed by regulation such as GAAP, these firms have an incentive to add as much noise as possible or otherwise avoid compliance. The assumption of precommitment is even more tenuous when there is no disclosure requirement. Firms with unfavorable demand signals in Cournot and unfavorable cost signals in Bertrand would seek to transmit the information voluntarily, since they are better off doing so. Conceivably, this could be achieved by detailed disclosures about future conditions (on markets, technology, input prices, etc.) in financial reports, announcements to the press, interviews with media journalists, and the like.

Bertrand With Cost Uncertainty

Contrary to the Cournot/demand case, the Bertrand/cost combination generates an equilibrium in which the disclosure region is significantly smaller. Furthermore, the disclosure region becomes smaller as the degree of substitutability increases. Define $Z_i^D(\gamma)$ and $Z_i^{ND}(\gamma)$ as the set of z_i values that firm i would choose to disclose and not to disclose, respectively. We again examine the convex combination of two polar cases: $\gamma=1$ and $\gamma \to 0$. By following the same procedure, we find the unique equilibria for the two cases, as follows.

Proposition 7: In Bertrand competition with cost uncertainty, the Bayesian-Nash equilibrium strategy of firm i in the *ex post* setting when firms receive private information for certain or with arbitrarily low probability are characterized by:

$$Z_i^D(1) = \begin{cases} \left[\bar{z}_i, \dfrac{t^2\bar{z}_i + 4a(2+t)(1-t)}{8-3t^2}\right], & \forall t<1, \\ \emptyset & t=1, \end{cases}$$

and

---

[30] Of course, it must be impossible for the *ex ante* commitments to constitute an equilibrium.

Table 1

	$\gamma\to 0$			$\gamma=1$		
t	*Lower Threshold*	*Upper Threshold*	*Area Percent*	*Lower Threshold*	*Upper Threshold*	*Area Percent*
0.4	0	3.830	49.99%	−2.887	3.768	99.80%
0.8	0	1.842	46.73%	−0.813	1.757	75.24%
0.95	0	0.557	21.12%	−0.071	0.545	23.54%
0.99	0	0.118	4.70%	−0.003	0.118	4.80%

$$\lim_{\gamma\to 0} Z_i^D(\gamma)=\begin{cases}\left[0,\dfrac{4a(2+t)(1-t)}{8-3t^2}\right], & \forall t<1,\\ \emptyset & t=1,\end{cases}$$

respectively, where $\bar{z}_i=E[z_i|z_i\in Z_i^{ND}(1)]$.

With the assumptions of $z_i\sim N(0,1)$ and $a=5$, for example, the disclosure region is computed as shown in table 1.

Table 1 shows that the disclosure region is smaller than in the Cournot/demand case, although it does increase with γ. When t is extremely small (e.g., $t=0.4$), it becomes virtually impossible to hide information as $\gamma\to 1$. For larger values of t, the disclosure region is quite small. In fact, it is striking how sensitive disclosure incentives are to the value of t.[31] As t increases, firms lose protection from product differentiation. The sensitivity to t is another manifestation of the cut-throat nature of Bertrand competition. Signals disclosed tend to be unfavorable (i.e., higher costs) because a firm with unfavorable information wants to transmit the information to induce price increases. Furthermore, it is also straightforward to check that firms would, *ex ante*, have an incentive to precommit to nondisclosure.[32]

IV. Summary and Discussion

In this article, I have considered the relationship between competition and voluntary disclosure behavior. Incentives to disclose depend upon the type of competition and private information. Those firms in Cournot competition facing cost uncertainty and those in Bertrand competition facing demand uncertainty are willing to commit to information disclosure and will support mandated disclosure of the specific information. Without precommitment to disclosure, firms that receive unfavorable information ex post would not want to disclose, but will succeed in concealing this information if and only if other firms cannot determine whether they have information. Otherwise, they end up disclosing information on a voluntary basis. Therefore, it appears that mandating disclosure is not necessary when firms are informed for sure.

---

[31] To see what sort of values t would take, recall that t is the cross-price elasticities between two products in Bertrand competition evaluated at the equilibrium. Cross-price elasticities of demand have been estimated, to take some examples, as follows: +0.44 for natural gas with respect to (the price of) fuel oil, +0.81 for butter with respect to margarine; +0.28 for beef with respect to pork. See Frank (1991). Substitutability between two brands of cereal, electronic appliances, etc. would be expected to be much higher.

[32] A proof is provided in the appendix.

Firms would precommit *ex ante* to nondisclosure of demand data under Cournot or cost data under Bertrand. Accordingly, if there are moves to mandate disclosure, these firms are not expected to support such requirements. The strongest resistance would come from firms in Bertrand competition regarding cost data, especially if rules do not leave room for noise.[33] Firms in Cournot competition with a high degree of product differentiation would also oppose mandating disclosure rules (about demand data), if flexibility allowed noncompliance, but would support stringent rules. In such a case, mandated disclosure rules substitute for a binding agreement, and both firms and consumers are better off because their interests are aligned. Also, both are better off with higher variation in outputs. This is not the case, however, when products are good substitutes. Although firms in Bertrand competition with cost uncertainty would also be worse off with mandatory disclosure, consumer surplus would be increased by such a policy.

Notwithstanding precommitted disclosure policies in place, firms with unfavorable signals have incentives to undermine such policies by disclosing information (when precommitment was to nondisclosure) or by providing only noisy data (when precommitment was to disclosure). Without a precommitted disclosure policy, firms are free to make voluntary disclosure. If a firm received a signal when other firms attach little likelihood to such an event, firms with favorable information would choose a disclosure policy that is identical to the one chosen *ex ante*, while firms with all but the most extreme unfavorable information would choose the opposite policy. When firms attach greater likelihood to the receipt of signals, however, it is in general more difficult to withhold information. Since nondisclosure is interpreted as having received extreme signals with high probability, firms would be forced to disclose more frequently.

In the Cournot/demand and Bertrand/cost combinations (where precommitment would have been to nondisclosure), firms with no precommitment would be compelled to disclose virtually all the time in the Cournot/demand case, but much less frequently in the Bertrand/cost case. In Bertrand competition with cost uncertainty, the frequency of disclosure and the range of values disclosed depend on the variability of the signal in relation to the market demand and the degree of substitutability. The higher the variance of the uncertainty parameter and the degree of substitutability, the less disclosure.

Not surprisingly, the implications of the analysis for social welfare are mixed. The welfare of consumers might be diametrically opposed to that of firms. The preference of firms is based on whether they are better off with or without disclosure (i.e., higher expected profits). Higher expected profits are achieved by choosing a disclosure strategy that raises expected equilibrium outputs and (net) prices. Higher expected outputs raise consumer surplus, but higher expected prices do not. An interesting result is that when products are not good substitutes, $(t<2\sqrt{2}-2)$, mandatory disclosure of demand data for Cournot firms enhances both firm profits and consumer surplus.

---

[33] An example of a requirement that left room for noise is FASB No. 33 on current costs. Since data on replacement costs were not required to be audited, it was easy to add noise. It is interesting to note that very little cost information is disclosed in financial reports in the United States, whereas Japanese firms typically provide manufacturing cost data broken down into categories of material, labor, and overhead.

Clearly, mandatory disclosure leads to a Pareto improvement and is socially desirable in this particular situation.

The assumption that disclosure is noiseless and credible may not be as innocuous as it seems. When private information concerns future prospects, credible disclosure may not be as easily achievable as it is when the information is about past (realized) performance, such as actual sales or total costs. As has been discussed in Okuno-Fujiwara et al. (1990), it may be more difficult to convince others that the disclosure is credible when a firm has unfavorable information. Favorable market demand may be supported by customer orders or well-documented market research data. Favorable cost conditions could be supported by technology specifications, procurement of raw materials at low prices, discovery of new methods, and the like. Unfavorable information, however, might require negative proof of the nonexistence of favorable conditions. To the extent that disclosure of certain types of information is more difficult, it is easier for firms to succeed in withholding this kind of information.

Another assumption, that the administrative cost of disclosure is negligible, permits a focus on a specific type of proprietary costs.[34] Firms are assumed to choose a disclosure strategy to maximize their expected profits, or "intrinsic value," in a duopolistic product market. Disclosure incentives may be made more complex by introducing additional markets, such as a financial market or another product market. Consider a firm that needs financing. Disclosure of favorable information to the financial market may increase firm valuation, but may compromise competitive strategy in a product market. If firms are playing Bertrand/demand or Cournot/cost games however, there are no conflicts. Disclosure does not involve a proprietary cost from competition. A firm with favorable information benefits from disclosure in both markets, while a firm with unfavorable information would try in vain to withhold the information from both markets. Hence, it is clear that full disclosure will ensue.

Conflicts arise in Cournot/demand and Bertrand/cost combinations.[35] In a Cournot game, a firm with information on high demand would want to conceal it to prevent the other firm from increasing output, yet would want to transmit good news to the financial market to increase its valuation. A firm with information on low demand would wish to withhold this information from the financial market, but disclose it to the rival to discourage overproduction. These conflicting motives are completely parallel in Bertrand competition with cost uncertainty. A low-cost firm would not want its rival to know, but would want the financial market to know the information to achieve more favorable financial valuation.

Another example is a duopolist in a product market who also operates in another product market. Action taken by the firm in one market might affect the equilibrium in the other market. Bulow et al. (1985) discuss a model in which disclosure in one market adversely affects the firm's overall profit from the other markets. This result depends on whether products are "strategic substitutes" "or complements" in the two markets.

[34] Administrative costs may not be negligible; they are, nevertheless, not interesting for analysis.

[35] A Cournot/demand case has been analyzed by Darrough and Stoughton (1990) and Wagenhofer (1990) in an entry game, and by Dontoh (1990) in a duopoly. In a duopoly setting, Gigler (1992) relaxes the assumption of truthful disclosure and establishes partition equilibria.

Since the results depend on the specifics of the model developed, it is worthwhile to examine the implications of some of the assumptions made here: (1) that signals received are independent, (2) that products are (imperfect) substitutes, and (3) that there are only two firms. The assumption of independence simplified the analysis and is easy to justify to the extent that the products are differentiated. Nevertheless, it might be unreasonable to assume totally uncorrelated signals (1) if signals are about a common value such as demand or (2) if they are about firm-specific costs but share similar technology and use similar inputs. An alternative scenario is positive correlation. Firms are then able to conjecture their rival's signal better by looking at their own private signals. To this extent, the incentive to share would diminish. At the same time, the benefit from withholding information would also diminish. The net effect is not obvious. Since perfect correlation negates any effects of disclosure, it is likely that the overall effect of disclosure would diminish as correlation increases. My conjecture is that the basic tradeoffs would remain the same.[36] In the *ex post* scenario, however, as correlation increases, there would be less divergence in the firms' preferences. It is possible that firms in an industry may lobby for a particular disclosure policy *ex ante*, then change their position completely once they receive signals that are highly (and positively) correlated.

The assumption that products are substitutes is obviously important. If they are independent, there is no issue of disclosure. If they are complements, firms help each other rather than divide the market. All the results are reversed.[37] Restricting our analysis to a duopoly does not affect the qualitative results. If the number of firms is large, however, other issues may emerge, such as whether a subgroup of firms wants to form a coalition.[38]

Appendix

Proof of Lemma

At the second stage, firm *i*'s best response is to choose Q_i so as to maximize expected profit, $E\Pi_i$, conditioned on the realization of x_i and its conjecture about firm *j*'s signal, given a disclosure (or lack thereof):

$$E\Pi_i = E_{x_j}\{Q_i[a+E(\Delta a)-Q_i-tE(Q_j)]\}. \tag{A1}$$

Since marginal costs are certain, $\sigma_c = n = 0$ and $\Delta c_1 = \Delta c_2 = 0$. To maximixe equation (A1), taking Q_j as given (since the conjectural variation in a Cournot game is zero), the partial derivative of $E\Pi_i$ with respect to Q_i is set to zero.

$$\frac{\partial E\Pi_i}{\partial Q_i} = a + E[\Delta a - btE(Q_j)] - 2Q_i = 0.$$

---

[36] Gal-Or (1985) examines a Cournot/demand case with a homogeneous good when private signals are correlated. The equilibrium found is again no disclosure, but equilibrium strategies are strictly Nash rather than dominant. Shapiro (1986) also finds that, when costs are uncertain but correlated, disclosure is optimal as long as correlation is positive and less than 1.

[37] See Vives (1984) for an analysis of a Cournot duopoly with demand uncertainty when products can be either substitutes or complements.

[38] See Kirby (1990) and Vives (1990).

The second-order condition is satisfied, since:

$$\frac{\partial^2 E\Pi}{\partial Q_i^2} = -2 < 0.$$

Hence,

$$Q_i = \frac{1}{2b}\{a + E[\Delta a - tE(Q_j)]\}. \tag{A2}$$

The posterior expected value of Δa is:

$$E(\Delta a) = \frac{1}{2}\left(\frac{\sigma}{m+\sigma+s_j}\hat{x}_j + \frac{\sigma}{m+\sigma}x_i\right). \tag{A3}$$

Given the message by firm j of $\hat{x}_j$, firm i updates $E(\Delta a)$ as a weighted average of j's message and its own private signal, with the weights being the precision of the message and the signal. Furthermore, posterior beliefs (conditional expectations) are:

$$E(x_j|\hat{x}_j) = \frac{m+\sigma}{m+\sigma+s_j}\hat{x}_j, \tag{A4}$$

$$E(x_i|\hat{x}_i) = \frac{m+\sigma}{m+\sigma+s_i}\hat{x}_i, \text{ and} \tag{A5}$$

$$E(\Delta a_i|x_i) = \frac{\sigma}{m+\sigma}x_i. \tag{A6}$$

To derive equation (2), one conjectures that Q_j is linear in information variables as:

$$Q_j = A_0^j + A_1^j\hat{x}_i + A_2^j\hat{x}_j + A_3^j x_j \qquad i \neq j = 1,2. \tag{A7}$$

Substituting equation (A7) into equation (A2), by using posteriors and making sure that Q_i is also linear in $\hat{x}_j$, $\hat{x}_i$ and x_i, yields the coefficients in equation (2). That is:

$$Q_i = \frac{1}{2b}\{a + E(\Delta a) - t[A_0^j + A_1^j\hat{x}_i + A_2^j\hat{x}_j + A_3^j E(x_j|\hat{x}_j)]\},$$

$$= \frac{1}{2b}\left\{a - btA_0^j - btA_1^j\hat{x}_i - \left[tA_2^j + tA_3^j\frac{(m+\sigma)}{m+\sigma+s_j} - \frac{\sigma}{2(m+\sigma+s_i)}\right]\hat{x}_j + \frac{\sigma}{2(m+\sigma)}x_i\right\},$$

$$= A_0^i + A_1^i\hat{x}_j + A_2^i\hat{x}_i + A_3^i x_i.$$

Solving for symmetrical A_k^i and A_k^j, $k = 0,1,2,3$ obtains equation (2).

Proof of Proposition 1

Given the second-stage quantity choice, expected profit for firm i at the first stage is:

$$E\Pi_i = E_{x_j,x_i}(Q_i^2)$$

$$= \frac{1}{(2+t)^2}\left[a^2 + \frac{\sigma^2}{4(m+\sigma+s_j)} - \frac{t(4+t)\sigma^2}{16(m+\sigma+s_i)} + \frac{\sigma^2(2+t)^2}{16(m+\sigma)}\right], \tag{A8}$$

when using the relations that $E(x_i) = E(x_j) = E(\hat{x}_j) = E(\hat{x}_i) = 0$. Taking the partial derivative of equation (A8) with respect to s_i yields:

$$\frac{\partial E\Pi_i}{\partial s_i} > 0.$$

Proof of Propositions 2 through 4

The derivation of these propositions is analogous to that for Proposition 1. The expected equilibrium profits are listed below for Bertand/demand (Bd), Cournot/cost (Cc), and Bertrand/cost (Bc) cases, respectively:

$$E\Pi_i^{Bd}=\frac{1}{1-t_2}E_{x_j,x_i}(P_i^2)$$

$$=\frac{(1-t)}{(1+t)(2-t)^2}\left[a^2+\frac{\sigma^2}{m+\sigma+s_j}+\frac{t\sigma^2}{2(m+\sigma+s_i)}+\frac{(2-t)^2\sigma^2}{16(m+\sigma)}\right].$$

$$E\Pi_i^{Cc}=\frac{1}{(2+t)^2}\left[a^2+\frac{t^2\sigma_c^2}{(2-t)^2(n+\sigma_c+v_j)}+\frac{t^2\sigma_c^2(8-t^2)}{4(2-t)^2(n+\sigma_c+v_i)}+\frac{(2+t)^2\sigma_c^2}{4(n+\sigma)}\right].$$

$$E\Pi_i^{Bc}=\frac{(1-t)}{(1+t)(2-t)^2}\left[a^2+\frac{t^2\sigma_c^2}{(2+t)^2(1-t)^2(n+\sigma_c+v_j)}\right.$$

$$\left.-\frac{t^2\sigma_c^2(8-3t^2)}{4(2+t)^2(1-t)^2(n+\sigma_c+v_i)}+\frac{(2-t)^2(\sigma_c+4n)}{4(1-t)^2(n+\sigma_c)}\right].$$

Proof of Proposition 5

The proof follows from the argument in the text.

Proof of Proposition 6

The remaining three cases are discussed here. Preferences are shown by looking at expected equilibrium prices and quantities.

1. *Bertrand/Demand Case.* Without any private signal or a message from firm 1, firm 2 will choose its price as:

$$P_2=\frac{a(1-t)}{2-t}.$$

With a private signal, x_1, firm 1 chooses its price, if it decides not to send any message, as:

$$P_1^{ND}=\frac{1-t}{2-t}\left[a+\frac{2-t}{2}E(\Delta a|x_1)\right].$$

If it decides to disclose, firm 1's price will be:

$$P_1^D=\frac{1-t}{2-t}\left[a+tE(\Delta a|\hat{x}_1)+\frac{2-t}{2}E(\Delta a|x_1)\right].$$

Thus, $P_1^D\geq P_1^{ND}$ for $x_1\geq 0$.

2. *Cournot/Cost Case.* Without firm 1's disclosure, firm 2 will choose $Q_2=a/[b(2+t)]$. Absent disclosure, firm 1 chooses:

$$Q_1^{ND}=\frac{1}{(2+t)}\left[a-\frac{2+t}{2}E(\Delta c|z_1)\right],$$

whereas with disclosure it will produce:

$$Q_1^D=\frac{1}{(2+t)}\left[a-\frac{2+t}{2}E(\Delta c|z_1)-\frac{t^2}{2(2+t)}E(\Delta c|\hat{z}_1)\right].$$

It is clear that $Q_1^D\geq Q_1^{ND}$ for $z_1\leq 0$.

3. *Bertrand/Cost Case.* Firm 1 chooses its (net) price without disclosure as:

$$P_1^{ND}=\frac{1-t}{2-t}\left[a-\frac{2-t}{1-t}E(\Delta c|z_1)\right],$$

while with disclosure, its price will be:

$$P_1^D=\frac{1-t}{2-t}\left[a+\frac{t^2}{2(2+t)(1-t)}E(\Delta c|\hat{z}_1)-\frac{2-t}{2(1-t)}E(\Delta c|z_1)\right].$$

Thus, $P_1^{ND}\geq P_1^D$ for $z_1\leq 0$.

Proof of Proposition 7

When $\gamma=1$, the expected profit of firm i is:

$$E\Pi_i=\frac{1}{(2-t)^2(1+t)}\left[a+\frac{t}{(2+t)(1-t)}\bar{z}_j+\frac{t^2}{2(2+t)(1-t)}(\bar{z}_i-z_i)-\frac{2-t^2}{(2+t)(1-t)}z_i\right]^2,$$

which gives,

$$E\Pi_i^{ND}-E\Pi_i^{D}=\frac{t^2}{(4-t^2)(1-t^2)}(\bar{z}_i-z_i)\left[\frac{t^2}{4(2+t)(1-t)}\bar{z}_i-a+\frac{8-3t^2}{4(2+t)(1-t)}z_i\right],$$

where $\bar{z}_i=E(z_i|z_i\in Z_i^{ND})$ and $\bar{z}=0$ (when $\gamma=0$ and $\bar{z}\neq 0$) (when $\gamma=1$). The difference in expected profits is nonpositive in the interval in the proposition. When $\gamma=0$, let $\bar{z}_i=0$ and solve for z_i.

To prove that firms prefer to precommit themselves to nondisclosure in Bertrand with cost uncertainty, let $\pi_i(\gamma)$ be the probability that $x_i\in X_i^{ND}(\gamma)$ for $i\neq j=1,2$. Then expected profit of *i* under a no-precommitment regime is [multiplied by $b(2-t)^2(1+t)$]:

$$\begin{aligned}E\Pi_i^{NC}=E\Bigg(\pi_j\left[a+\frac{t}{(2+t)(1-t)}\bar{z}_j-\frac{2-t^2}{(2+t)(1-t)}z_i\right]^2\\
+(1-\pi_j)E_{z_j}\left\{\left[a+\frac{t}{(2+t)(1-t)}\bar{z}_j-\frac{2-t^2}{(2+t)(1-t)}z_i\right]^2|z_j\in Z_j^D\right\}\Bigg)\\
+\pi_i\frac{8-3t^4}{4(2+t)^2(1-t)^2}\mathrm{Var}(z_i\,|\,z_i\in Z_j^{ND}).\end{aligned}$$

whereas expected profit under a precommitment (to no disclosure) regime is:

$$E\Pi_i^C=E\left\{\left[a-\frac{2-t^2}{(2+t)(1-t)}z_i\right]^2\right\}+\frac{8-3t^4}{4(2+t)^2(1-t)^2}\sigma_c^2.$$

This implies

$$\begin{aligned}E\Pi_i^C\geq E\Pi_i^{NC}\Leftrightarrow\frac{8-3t^4}{4(2+t)^2(1-t)^2}[\sigma_c^2-\pi_i\mathrm{Var}(z_i|z_i\in Z_i^{ND})]\\
\geq\frac{t^2}{(2+t)^2(1-t)^2}[\sigma_c^2-\pi_j\mathrm{Var}(z_j\,|\,z_j\in Z_j^{ND})],\\
=\frac{t^2}{(2+t)^2(1-t)^2}[\sigma_c^2-\pi_i\mathrm{Var}(z_i|z_i\in Z_i^{ND})],\end{aligned}$$

where the last equality follows from symmetry. This boils down to a condition $8-3t^4\geq 4t^2$, which holds for any value of $0\leq t\leq 1$.

References

Bulow, J. I., J. D. Geanakoplos, and P. D. Klemperer. 1985. Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political Economy* 93: 488–511.

Clarke, R. N. 1983. Collusion and incentives for information sharing. *Bell Journal of Economics* 14 (Autumn): 383–94.

Clinch, G., and R. E. Verrecchia. 1991. Information sharing in oligopoly with no precommitment. Working paper, Wharton School, University of Pennsylvania.

Darrough, M. N., and N. M. Stoughton. 1990. Financial disclosure policy in an entry game. *Journal of Accounting & Economics* 12 (January): 219–43.

Dontoh, A 1990. Voluntary disclosure. *Journal of Accounting, Auditing & Finance* 4 (Fall): 480–511.

Dye, R. 1986. Proprietary and nonproprietary disclosures. *Journal of Business* 59: 331–66.

Feltham, G. A., and J. Z. Xie. 1992. Voluntary financial disclosure in an entry game with continua of types. *Contemporary Accounting Research* 9 (Fall): 46–80.

——, F. B. Gigler, and J. S. Hughes. 1992. The effects of line of business reporting on competition in oligopoly setting. *Contemporary Accounting Research* 9 (Fall): 1–23.

Foster, G. 1986. Financial Statement Analysis. Englewood Cliffs, NJ: Prentice-Hall.

Frank, R. H. 1991. *Microeconomics and Behavior*. New York: McGraw-Hill.

Fried, D. 1984. Incentives for information production and disclosure in a duopolistic environment. *Quarterly Journal of Economics* 99 (May): 367–81.

Gal-Or, E. 1985. Information sharing in oligopoly. *Econometrica* 53 (March): 329–43.

——. 1986. Information transmission—Cournot and Bertrand equilibria. *Review of Economic Studies* 53: 85–92.

Gertner, R., R. Gibbons, and D. Sharfstein. 1988. Simultaneous signalling to the capital and product markets. *Rand Journal of Economics* 19: 173–90.

Gigler, F. B. 1992. Self-enforcing public disclosures. Working paper, University of Chicago.

Grossman, S. 1981. The information role of warranties private disclosure about product quality. *Journal of Law and Economics* 24: 461–83.

Hughes, J. S., and J. Kao. 1991. Economic implication of alternative accounting rules for research and development costs. *Contemporary Accounting Research* 8 (Fall): 152–69.

Jung, W. O., and Y. K. Kwon. 1988. Disclosures when the market is unsure of information endowment of managers. *Journal of Accounting Research* 26: 146–53.

Kirby, A. 1988. Trade associations as information exchange mechanisms. *Rand Journal of Economics* 19 (Spring): 138–46.

——. 1990. Optimal information exchange. Working paper, University of California, Berkeley.

Klemperer, P., and M. Meyer. 1986. Price competition vs. quantity competition: The role of uncertainty. *Rand Journal of Economics* 17 (Winter): 618–38.

Milgrom, P. 1981. Good news and bad news: Representation theorems and application. *Bell Journal of Economics* 12: 380–91.

Novshek, W., and H. Sonnenschein. 1982. Fulfilled expectations Cournot duopoly with information acquisition and release. *Bell Journal of Economics* 13: 214–18.

Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura. 1990. Strategic information revelation. *Review of Economic Studies* 57: 25–47.

Sankar, M. 1992. Corporate voluntary disclosure of predecision information. Ph.D. dissertation. University of British Columbia, Vancouver.

Shapiro, C. 1986. Exchange of cost information in oligopoly. *Review of Economic Studies* 53: 433–46.

Verrecchia, R. 1983. Discretionary disclosure. *Journal of Accounting & Economics* 5 (December): 179–94.

——. 1990a. Endogenous proprietary costs through firm interdependence. *Journal of Accounting & Economics* 12 (January): 245–50.

——. 1990b. Information quality and discretionary disclosure. *Journal of Accounting & Economics* 12 (March): 365–80.

Vives, X. 1984. Duopoly information equilibrium: Cournot and Bertrand. *Journal of Economic Theory* 34: 71–94.

——. 1990. Trade association disclosure rules, incentives to share information, and welfare. *Rand Journal of Economics* 21 (Autumn): 409-30.

Wagenhofer, A. 1990. Voluntary disclosure with a strategic opponent. *Journal of Accounting & Economics* 12 (March): 341–63.

A FORUM ON
COST DRIVERS

1. Yair M. Babad and Bala V. Balachandran
   Cost Drivers Optimization in Activity-Based Costing

2. Rajiv D. Banker and Holly H. Johnston
   An Empirical Study of Cost Drivers in the U.S. Airline Industry

3. Srikant M. Datar, Sunder Kekre, Tridas Mukhopadhyay, and Kannan Srinivasan
   Simultaneous Estimation of Cost Drivers

4. Nicholas Dopuch
   A Perspective on Cost Drivers

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 563–575

Cost Driver Optimization in Activity-Based Costing

Yair M. Babad
University of Illinois at Chicago
Bala V. Balachandran
Northwestern University

SYNOPSIS AND INTRODUCTION: The goal of any cost management system is to provide relevant and timely information to management. This information supports better management of corporate resources in production of products or provision of services, and improves competitiveness in terms of costs, quality, and profitability. In this context, a cost management system can also be viewed as a planning and control management system (Berliner and Brimson 1988). Cooper (1988a, 1988b, 1989a, 1989b) provided a comprehensive discussion of activity-based costing (ABC), following the pioneering work of Kaplan (1983, 1984). Extension of ABC to the service industry was provided by Rotch (1990).

ABC has also been extended into activity-based management (ABM) to include other considerations, such as customer profitability, manpower utilization, distribution channels, and other management issues. Thus, ABC is the information system that reveals the cost and profitability structure of products and services in an organization, while ABM describes the actions taken to improve quality and reduce costs and cycle time, once information about activities' costs is known. In this article, ABC is used as a generic concept without any loss of generality.

An ABC system achieves improved accuracy in estimation of costs by using *multiple* cost drivers to trace the cost of activities to the products associated with the resources consumed by those activities. In this respect, a *cost driver* is an event, associated with an activity, that results in the consumption of firm's resources. Since the number of events performed in a firm is often vast, it may not be cost-effective to use a distinct and different cost driver for each activity. Thus, many activities may be grouped into a single driver to trace the costs of all the grouped activities for a product or

The authors express their sincere appreciation to the three referees and to Professor Sri Sridhar for their helpful comments. The second author expresses his appreciation to Ernst & Young and the Northwestern Accounting Research Center for their financial support.

Submitted January 1991.
Accepted May 1993.

service. For instance, each setup may be associated with a single cost driver that accounts for moving, grouping, sequencing, and segmenting. At the same time, there may be other competing cost drivers, such as setup hours, better correlated with the resources consumed by these grouped activities. In activity-based costing, these different cost drivers are not necessarily all proportional to unit volume, in contrast to traditional volume-based cost systems (Kaplan 1988, O'Guinn 1990, Dewan and Magee 1992).

"The art of designing an ABC system can be viewed as making two separate but interrelated decisions about the *number of cost drivers* needed and *which cost drivers to use*. These decisions are interrelated because the type of cost drivers selected changes the number of drivers required to achieve a desired level of accuracy" (Cooper 1989a, 1989b). In this article, we provide an optimization model that balances savings in information processing costs with loss of accuracy. We show how to determine the number of drivers, and identify the representative cost drivers. The model is formulated as an integer program and is solved efficiently by using a composite greedy algorithm.

Key Words: *Activity-based costing, Cost drivers, Optimization, Cost benefit trade off, Cost management.*

In section I, we define the variables used in our model. In section II, the accuracy of a system of multiple cost drivers is defined, and the effect of combinations of cost drivers on accuracy is analyzed. In section III, the optimization model is analytically presented, and a solution algorithm is provided. In section IV, we provide a numerical illustration, then summarize our findings and present suggestions for extensions in section V.

I. The Activity-Based Cost Model

Activity-based costing has recently emerged as a superior method for obtaining more accurate product costs and thus, better factory management. In addition, ABC provides insights for the analysis of customer profitability, distribution channels, vendors, brands, regions, and other areas that directly affect the profitability of a firm. Because ABC links performance of particular activities (frequency) with the demands these make on the organization's resources, the revenues from products or services can be directly matched with costs as resources consumed (Cooper and Kaplan 1992). Ideally, to derive *true* costs, one first considers *every* activity of the company, where all activities are mutually exclusive and collectively exhaustive, then identifies the unique cost driver associated with each activity to develop a per activity cost. Finally, the costs of all activities consumed by a product or process are multiplied by their frequency and are accumulated. This is known as the "bill of activities," and is similar to the "bill of materials." For detailed definitions of ABC terms, see O'Guinn (1991).

This approach is in contrast with the traditional cost-accounting system that maintains one big overhead pool with a *single* activity rate, for example, one based on direct labor, with the number of direct labor hours (or dollars) consumed as the cost driver. The single activity is easy for data gathering, storing, and processing, but may be highly erroneous and dysfunctional. Use of all activities in minute detail, in contrast, may be more accurate but enormously costly to collect, store, and process. Even if data for some cost drivers are already being collected by existing systems, the transfer and merging of this data into the total product cost may be a costly and complex process. Furthermore, errors of data collection, reporting, and estimation of sources, and possible inconsistencies between different cost drivers that are generated by different systems can also create errors in product and process costing. Finally, cost drivers are also chosen for planning and control; if too many drivers are chosen, planning and controlling the production cost for each would be neither effective nor efficient. Thus, on a cost-benefit basis, a balance must be sought between accuracy benefits and costs of data collection, storage, and processing. Only those drivers that are related to the bulk of costs need to be identified and tracked, and organizations such as GM's Saturn, Kraft, and John Deere tend to restrict the number of cost drivers actually used to between six and nine, despite the availability of many more drivers.

To determine the optimal balance, one must find the optimal number of cost drivers that discriminates and captures most of the incurred costs, and identify a priority order that specifies which low-priority and relatively insignificant activities will be combined to save costs without sacrificing much accuracy. We mathematically prove (in Lemma 1 below) the correctness of the intuitive notion that if any two or more activities are perfectly correlated, there is no benefit of improved accuracy in keeping them separate. Thus, they can be combined into one major activity with an associated cost driver. However, combining significant activities that are *not* perfectly correlated can lead to some loss of accuracy in product and process costing. In the model given below, we want to optimize this cost-benefit tradeoff through the balance of loss of accuracy with the reduction in number of cost drivers by combining activities.

Once a desired accuracy level is selected, according to priorities, economy, profit margin squeeze, and the competitive environment, the model identifies the minimum number and order of activities and drivers to achieve that accuracy. The manager can perform a sensitivity analysis and see the effect of changing accuracy levels. The key to our analysis is the use of *relative*, rather than actual or absolute, frequency (or volume of use) of a cost driver associated with a product or process. Thus, this methodology becomes independent of the actual volume, historical average, or projected use of the activity. Rather, a probability interpretation is possible, as well as comparative analysis, similar to the use of statistical properties of minimum variance available in the methodology of least squares.

For our analysis, let I be the considered set of products or cost objects, and let J be the set of cost drivers associated with the production activities (hereafter we will use cost drivers and activities interchangeably). A subscript i will be used for a product $(i \epsilon I)$, and a subscript j is used for a cost driver $(j \epsilon J)$. Let:

$\bar{V}_{ij}$ = volume, or the actual absolute frequency, of use of cost driver j by product i,
D_j = total dollar expenses of activities associated with cost driver j, *and*
U_i = total cost of product i, after taking into account the contribution of all the cost drivers:

$$U_i = \sum_j D_j \frac{\bar{V}_{ij}}{\sum_i \bar{V}_{ij}}.$$

Note that in this expression, the actual absolute frequencies are effectively converted (through the division by the total volume for each cost driver) into relative frequencies of use. Thus, it is appropriate to define:

V_{ij}=relative frequency of use of cost driver j by product i, where:

$$V_{ij}=\frac{\bar{V}_{ij}}{\sum_i \bar{V}_{ij}}.$$

The total cost of product i is then:

$$U_i=\sum_j D_j V_{ij}, \tag{1}$$

and the total cost for the whole production is:

$$U=\sum_i U_i.$$

It is possible that unused or idle capacity exists for some activities. Our relative frequencies may associate excess capacity costs with products, in violation of ABC theory (Cooper and Kaplan 1992). To overcome this problem, we use a fictitious "null" product to identify and accumulate the unused capacity costs.

The transformation given above eliminates the actual volume of use of a driver and replaces it with the relative frequency (simply frequency in the sequel) for each product, where these frequencies sum to one for each cost driver. Thus, the selection and determination of cost drivers become independent of the actual, historical, or projected volume of use, and relative frequency of use is instead the parameter of interest.

Our goal is to minimize the cost of accumulating, carrying, and managing the information required by the cost drivers, without sacrificing too much accuracy. This may require the combination of some cost drivers. In this case, the frequencies of the remaining cost driver include the contributions of the replaced cost drivers. Such a combination may include both planned cost drivers that are contemplated for use, as well as cost drivers for which data is already available from other systems.

II. The Accuracy of Combined Cost Drivers

To construct a model for optimal selection of cost drivers, we must first determine how to measure the accuracy of the ABC system and any sacrifice in accuracy when cost drivers are combined. Let U_i^{km} be the cost of product i, after taking into account the contribution of all the cost drivers and combining cost drivers k and m into cost driver m. Then,

$$U_i^{km}=\sum_{j\neq k,m} D_j\bullet V_{ij}+(D_k+D_m)\bullet V_{im}, \quad \text{and using equation (1),}$$

$$=U_i+D_k\bullet V_{im}-D_k\bullet V_{ik}=U_i+D_k\bullet(V_{im}-V_{ik}), \text{ and}$$

$$U_i-U_i^{km}=D_k\bullet(V_{ik}-V_{im}). \tag{2}$$

Equation (2) expresses the *accuracy* of the combination km with respect to product

i and enables us to evaluate the effect of the combination on the cost of each individual product. In a particular case, described below, a reduction in the number of cost drivers, and thus, also in the total cost of maintaining the ABC system, can be accomplished without any loss of accuracy. To this end, define two cost drivers as *perfectly correlated*, when they have the same frequency values (i.e., if $V_{ik}=V_{im}$ for cost drivers k and m and all products $i\epsilon I$).

Lemma 1: Perfectly correlated cost drivers can be combined, that is, can be represented by one of the correlated drivers, after dropping the other correlated drivers, without any loss of accuracy in the products' costs.[1]

Proof: It is sufficient to consider a pair of perfectly correlated cost drivers, say k and m. Then, for an arbitrary product $i\epsilon I$:

$$U_i=\sum_{j} D_j\bullet V_{ij}=\sum_{j\neq k,m} D_j\bullet V_{ij}+D_k\bullet V_{ik}+D_m+V_{im},$$

and since k and m are perfectly correlated,

$$=\sum_{j\neq k,m} D_j\bullet V_{ij}+(D_k+D_m)\bullet V_{ik}.$$

Thus, cost driver k can replace and eliminate cost driver m.

When two cost drivers are perfectly correlated, the decision about which to keep and which to eliminate will usually take into account the cost and difficulty of maintaining and processing each cost driver. Note that this lemma mathematically justifies the intuitive notion that, with a single product, all expenses are applied to the same product and there is no need for an activity-based costing system. Indeed, in this case, $V_{1j}\equiv 1$ for all the cost drivers (j); thus, all cost drivers are perfectly correlated and activity-based costing does not add anything (beyond a traditional cost system).

From Lemma 1, and without loss of generality, assume that all the perfectly correlated cost drivers have been combined, and no perfectly correlated cost drivers remain in the set J. Note that the nearer the accuracy is to zero, the better the match (costwise) between the original set of cost drivers and the combined one. Furthermore, since, $\Sigma_i V_{ik}=\Sigma_i V_{im}=1$, and since the cost drivers are not perfectly correlated (so that for at least one product x we have $V_{xk}\neq V_{xm}$), we know that the sign of $V_{ik}-V_{im}$ changes with i so that there are products x and y such that $V_{xk}-V_{xm}<0<V_{yk}-V_{ym}$. In other words, equation (2) implies that, once cost drivers are combined, the costs of certain products will be overestimated, and that of others will be underestimated. At the same time, the total cost of production will not change, since if we sum equation (2) over all the products $i\epsilon I$, we will get:

$$U-\sum_{i} U_i^{km}=\sum_{i}(U_i-U_i^{km})=D_k\bullet\left(\sum_{i} V_{ik}-\sum_{i} V_{im}\right)=0.$$

These observations suggest that the *squared distance* of the products' costs, $\Sigma_i(U_i-U_i^{km})^2$, will be a better measure than the accuracy for the goodness of fit of a combination of cost drivers. However, when we consider the analogous statistical reasoning, we find that a nonsquared distance, which naturally has the dimension of dollars, may be more appropriate.

---

[1] Feltham (1977) has derived certain related results in a framework for "decisions under uncertainty." Our results are for a different context in a world of certainty.

As stated, the goal of minimizing the cost of accumulating, carrying, and managing the information required by the cost drivers without sacrificing too much accuracy, often requires that several cost drivers be combined. The need to consider all the feasible combinations of cost drivers then complicates any solution algorithm; Lemma 2, however, assures us that it is sufficient to consider only pairwise combinations of cost drivers.

Multiple cost drivers can be combined in various ways. They can be combined pairwise (i.e., cost drivers k_1 and m_1 are combined into cost driver m_1, k_2 and m_2 are combined into $m_2,\ldots,k_n$ and m_n are combined into m_n); so that the related cost for product *i is denoted as* $U_i^{(k1,m1),\ldots,(kn,m)}$. But they can all also be combined into a single cost driver (e.g., cost drivers $k_1,\ldots,k_n$ and m are all combined into cost driver m, which is denoted as $U_i^{(k1,\ldots,kn,m)}$, where the consumptions of $k_1,\ldots,k_n$ and m are summed together and the frequencies of cost driver m are used. These two ways are called *pairwise* and *multiple* combinations of cost drivers.

Lemma 2: A multiple combination of cost drivers can be replaced by a set of pairwise combinations of the same cost drivers, without loss of accuracy in the cost of the products.

Proof: Let $k_1,\ldots,k_n$ be the cost drivers that are combined into m. Then:

$$U_i^{(k1,\ldots,kn,m)} = \sum_{j\neq k1,\ldots,kn,m} D_j \cdot V_{ij} + (D_{k1},\ldots,D_{kn}+D_m) \cdot V_{im},$$

$$= \sum_j D_j \cdot V_{ij} + \sum_{x=1}^{n} (D_{kx} \cdot V_{im} - D_{kx} \cdot V_{ikx}), \text{ and}$$

$$= U_i + \sum_{x=1}^{n} D_{kx} \cdot (V_{im} - V_{ikx}).$$

Thus,

$$U_i - U_i^{(k1,\ldots,kn,m)} = \sum_{x=1}^{n} D_{kx} \cdot (V_{ikx} - V_{im}) \qquad \text{with equation (2) and}$$

$$= \sum_{x=1}^{n} (U_i - U_i^{kx,m}) = U_i - U_i^{(k1,m),\ldots,(kn,m)}.$$

III. Optimal Selection of Cost Drivers

The second element needed for the optimal selection of cost drivers is the cost associated with accumulating, carrying, and managing the information required by the cost drivers. Let C_j be the cost associated with cost driver j; then the full information cost is $\Sigma_j C_j$. The use of the combination (km) saves the need to collect, manage, verify, and use the information of cost driver k, and thus results in a savings of C_k. This savings, as noted before, may be worthwhile even for cost drivers for which cost data are already available.

Furthermore, not all products are "created equal"; some have more strategic or operational importance to the organization. Thus, we assign a weight w_i to product i, which is applied to its distance. Then, the term $(U_i - U_i^{km})^2 = [D_k \cdot (V_{i,k} - V_{i,m})]^2$ is replaced by $w_i \cdot [D_k \cdot (V_{i,k} - V_{i,m})]^2$.

The model maximizes the total weighted savings, that is, information costs savings minus the weighted distance costs. Following Lemma 2, which enables us to concentrate on pairwise combinations, we then have:

$$\underset{\text{all pairs } (k,m)}{\text{Max}} \quad \sum_{\text{all pairs } (k,m)} X_{km} \cdot \left\{ C_k - \text{SQRT} \left[\sum_i w_i \cdot (U_i - U_i^{k,m})^2 \right] \right\}$$

$$= \sum_{\text{all pairs } (k,m)} X_{km} \cdot \left\{ C_k - D_k \cdot \text{SQRT} \left[\sum_i w_i \cdot (V_{i,k} - V_{i,m})^2 \right] \right\}$$

where:

$$X_{km} = 1, \text{ if } (k,m)$$

is in the optimal selection, zero otherwise. Note that we use the square root of the squared distance, in a way similar to the determination of the standard deviation in statistics.

Practical considerations, such as processing and storage capacity or mangement's ability to comprehend and effectively use a large number of cost drivers, may limit the number of cost drivers to be used. If N is such a limit, the maximization model is then constrained by:

$$\sum_{\text{all pairs } (k,m)} X_{km} \leq N.$$

The model naturally requires that estimates of the parameters be available for all the cost drivers of interest, and it will be sensitive to the accuracy of these estimates; however, it does not require that these cost drivers already be in operation. Inaccuracies in the estimates can also be counterbalanced by an appropriate choice of the weights, as noted below. Further, although the savings that the model provides are attributable to the nonuse of some cost drivers for product and process management, these cost drivers need not be entirely eliminated. They may still be used for other purposes, even though their data are not assimilated, verified, merged, and used by the ABC system.

Note that the weights play a crucial role in the decision process, as they reflect the priorities of management and can be used to counterbalance inaccurate parameter estimates. The model, consequently, provides a way to quantify the options open to management, and optimizes the decision process so as to best reflect management's priorities. In our algorithm, the weight factor w_i for product i may also have a strategic effect. Often there are some vital products that create demand for other products, have high volume, and/or provide high profits. Such products would be given a (relatively) lower strategic weight w_is, to reflect the company's priorities among products. (Because the weights, and accuracies, are subtracted in the objective function of the model, the more preferred products should get smaller weights.) This is similar to the target-costing approach, where the price determines the affordable target cost. Alternatively, when such distinctions between products cannot be justified, one can set all the weights to be equal. To relect the emphasis the decision maker places on information cost savings relative to accuracy, one may then use high or low uniform weights. This provides support for the use of the model in planning an ABC system.

The information requirements for the model are relatively high, and the decision maker has to identify the potential cost drivers, their estimated usage costs, and management's priorities. As noted, estimates may be used in lieu of exact values. Thus, the model is particularly useful in the planning phases of an ABC system. Indeed, one should invest in such information to analyze the interplay of these parameters and their sensitivities. This way, the ongoing needs of informational costs, which usually are dominant in the long run, can be minimized before it is too late. System designers should think about the cost and benefit of various alternatives, and our model is a convenient tool for such analysis. The model can, of course, be used for "operational tuning" of an ABC system as more information is gathered during its actual operation.

The model itself is clearly an integer-programming problem, and thus, requires some computational effort for its solution. It is also a (zero, one) knapsack problem, with a knapsack of size N, a finite set of $J \cdot (J-1)$ "items" (i.e., pairs), each with a "size" of one (since we selected $x_{km}=1$), and with "value" $C_k - D_k \cdot \text{SQRT}[\Sigma_i w_i \cdot (V_{i,k} - V_{i,m})^2]$ for the item (km). In this problem, the objective is to place as many items as possible in the knapsack, so that their total value will be maximized. Also note that the model solves concurrently for both the number of pairs to be combined and the selection of the pairs themselves.

Integer-programming problems are usually hard to solve computationally; however, because of its particular structure, efficient optimal algorithms for the solution of the knapsack problem are available (Martello et al. 1987; Martello and Toth 1988). Here we will present a useful heuristic approximation, the composite greedy algorithm of Garey and Johnson (1979), as applied to our problem.

Step 1: Sort the cost driver pairs in descending order according to their knapsack "values" (in general these should be divided by the "sizes," which equal one in this case). The result is a list of ordered pairs of cost drivers. Set the "stack" of selected driver pairs to be combined to null.

Step 2: Add the first pair in the list to the stack. If this is the pair (k,m), that is the "source" driver k is combined into the "target" driver m, then:

2.1 eliminate from the list all other pairs with a source k, since this source has already been combined with some other target driver,

2.2 eliminate from the list all pairs with a source m, since this driver has already been chosen as a target for a combination of drivers.

2.3 If the stack already contains a pair (m,n), then m has already been selected to be combined into n. Therefore, replace the new pair on the stack, (k,m), with (k, n).

Repeat step 2 until a sufficient number of pairs are selected or until the list is exhausted.

Step 3: Compute the "value" for the selected solution and compare it to the "value" of the first pair on the original ordered list. If the value for the first pair exceeds that for the selected solution, ignore the solution and use instead the single pair that was the first on the list.

The major benefit of this approach is that the composite greedy algorithm assures a "value" close to the optimal "value" with relatively minimal computational effort. Although the algorithm does not assure the optimal solution, this is not usually a major drawback since no algorithm can be more accurate than the provided parameters and since many of these are just estimates.

Another benefit of this approach is that it provides an order of pairs merging, and thus their priority. The first step of the algorithm assures that the pairs with the highest value will be first selected. Thus, if the decision maker has some other constraints on the number of cost drivers to be used, aside from the desired level of accuracy, the algorithm can be stopped after the desired number of pairs have been selected.

IV. Example of Optimal Selection of Cost Drivers

To illustrate the concepts and algorithms, we adapted the example in Exhibit 2 of Cooper (1988a) for our table 1. In panel A, the input consumption patterns and dollar value data are displayed. A traditional cost system, which uses direct labor hours for overhead allocation, reports the product unit costs, as given in panel B. Note that under this traditional cost system, products P1 and P2 have the same unit cost, as have P3 and P4. An activity-based costing system utilizes the frequencies of the cost drivers' usage by products, as computed in panel C, and reports very different unit costs (panel D). It is evident that the traditional cost system leads to large inaccuracies in determining product costs, which may lead to erroneous decisions.

When Lemma 1 is applied to the data in panel C, it is immediately apparent that Material (Dollar) and Direct Labor Hours may be combined, say to Direct Labor Hours. Similarly, the Number of Orders and the Number of Times Handled may be combined, (e.g., into Number of Orders). The resulting set of input frequencies is shown in panel E. It is evident that, as Lemma 1 proves, the reduction in cost drivers is achieved without loss of accuracy, since the total overhead costs per product are the same.

Our algorithm requires that the related accuracy be determined for each product and pair of cost drivers. The nonweighted accuracies, $U_i - U_i^{km}$, are thus given in panel F of table 1, which also shows the nonweighted distances, SQRT$[\Sigma_i(U_i - U_i^{km})^2]$, as well as weighted distances (with weights as given at the top of the panel).

Note that the accuracy for any given product and pair of cost drivers is the difference between (1) product cost when all the cost drivers are used and (2) the cost after the combination is applied. Thus, to get the modified product's cost, one has to subtract the accuracy from the original cost. Also note the weights for the various products. As mentioned earlier, the lower the weight for a particular product, relative to other weights, the higher the importance of this product to the decision maker. In our example, P2 and P3 clearly are much more important than P1 and P4.

We assume that the costs of information accumulation, retrieval, and processing for the five cost drivers are as given at the end of panel E. Using these costs, we applied the algorithm to the weighted and the nonweighted data; the results are summarized in panel G. In the weighted case, (products 2 and 3 are considered more important than products 1 and 4), the algorithm recommended combining all the cost drivers into cost driver 2, machine hours. It is interesting to note the order, and priority, of the combinations: first cost driver 3 was combined with 2, then 4, 1, and finally 5. Note also how accurate the resulting costs for products 2 and 3, are relative to the costs for products 1 and 4. Finally, note that the algorithm produced costs that are uniformly superior in terms of accuracy to those of the traditional cost system based on direct labor hours.

Table 1
Product Costing Data for Company A

Panel A. Annual Input Consumption Patterns and Dollar Value by Product:

Product	*Quantity*	*Material (Dollar)*	*Direct Labor Hours*	*Machine Hours*	*Setups*	*Orders*	*Times Handled*	*Parts*	*Total Overhead Costs*
P1	10	60	5	10	1	1	1	1	
P2	100	600	50	40	2	3	3	1	
P3	10	180	15	20	1	1	1	1	
P4	100	1,800	150	100	4	3	3	1	
Units Consumed		2,640	220	170	8	8	8	4	
Dollar Value		$ 264	$2,200	$3,400	$960	$1,000	$200	$2,000	$10,024

Panel B. Annual Overhead Costs Reported by the Cost System:

Product	*Direct Labor Hours*	*Cost Trace = Total Overhead*	*Quantity*	*Reported Unit Cost*
P1	5	$ 227.82	10	$22.78
P2	50	2,278.18	100	22.78
P3	15	683.45	10	68.35
P4	150	6,834.55	100	68.35

Panel C. Annual Input Frequencies and Dollar Value by Product:

Product	*Quantity*	*Material (Dollar)*	*Direct Labor Hours*	*Machine Hours*	*Setups*	*Orders*	*Times Handled*	*Parts*	*Total Overhead Costs*
P1	10	0.0227	0.0227	0.0588	0.1250	0.1250	0.1250	0.2500	$ 1,026
P2	100	0.2273	0.2273	0.2353	0.2500	0.3750	0.3750	0.2500	2,550
P3	10	0.0682	0.0682	0.1176	0.1250	0.1250	0.1250	0.2500	1,338
P4	100	0.6818	0.6818	0.5882	0.5000	0.3750	0.3750	0.2500	5,110
Dollar Value		$ 264	$2,200	$3,400	$ 960	$1,000	$ 200	$2,000	$10,024

Panel D. Annual Overhead Costs Reported by the Activity-Based Costing System:

Product	*Cost Trace = Total Overhead*	*Quantity Per Year*	*Reported Unit Cost*
P1	$1,026.00	10	$102.60
P2	2,550.00	100	25.50
P3	1,338.00	10	133.80
P4	5,110.00	100	51.10

Continued

The costs and accuracies when the algorithm was applied to nonweighted products are presented at the end of panel G. In this case, the algorithm suggested that the largest benefit will accrue from the combination of cost driver 1 with 2. The algorithm then

Table 1—*Continued*

Panel E. Annual Input Frequencies and Dollar Value by Product, Reduced by Lemma 1:

Product	*Quantity*	*Cost Driver 1 Direct (Labor Hours)*	*Cost Driver 2 (Machine Hours)*	*Cost Driver 3 (Setups)*	*Cost Driver 4 (Orders)*	*Times Handled*	*Cost Driver 5 (Parts)*	*Total Overhead Costs*
P1	10	0.0227	0.0588	0.1250	0.1250		0.2500	$ 1,026
P2	100	0.2273	0.2353	0.2500	0.3750		0.2500	2,550
P3	10	0.0682	0.1176	0.1250	0.1250		0.2500	1.338
P4	100	0.6818	0.5882	0.5000	0.3750		0.2500	5,110
Dollar Value		$2,464	$3,400	$ 960	$1,200		$2,000	$10,024
Information Cost		2,500	1,500	2,000	2,000		2,500	

Panel F. Accuracies and Distances for Various Single Pair Combinations:

Combined Cost Drivers	*Accuracy for P1 (Weighted at 50)*	*Accuracy for P2 (Weighted at 1)*	*Accuracy for P3 (Weighted at 1)*	*Accuracy for P4 (Weighted at 50)*	*Nonweighted Distance*	*Weighted Distance*
(12)	−88.94	−19.76	−121.88	230.59	276.27	1751.95
(13)	−252.00	−56.00	−140.00	448.00	535.67	3637.74
(14)	−252.00	−364.00	−140.00	756.00	887.21	5648.37
(15)	−560.00	−56.00	−448.00	1064.00	1284.34	8514.03
(21)	122.73	27.27	168.18	−318.18	381.22	2417.46
(23)	−225.00	−50.00	−25.00	300.00	379.14	2652.24
(24)	−225.00	−475.00	−25.00	725.00	895.82	5388.76
(25)	−650.00	−50.00	−450.00	1150.00	1396.42	9351.74
(31)	98.18	21.82	54.55	−174.55	208.70	1417.30
(32)	63.53	14.12	7.06	−84.71	107.05	748.87
(34)	0.00	−120.00	0.00	120.00	169.71	856.97
(35)	−120.00	0.00	−120.00	240.00	293.94	1901.16
(41)	122.73	177.27	68.18	−368.18	432.08	2750.83
(42)	79.41	167.65	8.82	−255.88	316.17	1901.92
(43)	0.00	150.00	0.00	−150.00	212.13	1071.21
(45)	−150.00	150.00	−150.00	150.00	300.00	1514.93
(51)	454.55	45.45	363.64	−863.64	1042.49	6910.74
(52)	382.35	29.41	264.71	−676.47	821.43	5501.02
(53)	250.00	0.00	250.00	−500.00	612.37	3960.74
(54)	250.00	−250.00	250.00	−250.00	500.00	2524.88

Continued

combined cost driver 5 with 4, and then cost driver 3 with 2. Thus, the two cost drivers that should be used are 2, machine hours, and 4, number of orders. Note that the absolute magnitudes of the accuracies for the four products are much more uniform than in the weighted or traditional cost systems. Also note that the traditional cost driver of direct labor hours was not selected in either the weighted or nonweighted cases. Finally, note that the costs in the weighted and nonweighted cases are quite similar, but differ considerably from those in the traditional cost system.

Table 1—*Continued*

Panel G. Summary of Costs, Savings, and Distances After Algorithm is Applied—Product and Annual Quantity:

Product: *Annual Quantity:*	*P1* 10	*P2* 100	*P3* 10	*P4* 100	*Total* *Cost*
All Drivers					
Total Overhead	$1,026.00	$2,550.00	$1,338.00	$5,110.00	$10,024.00
Cost Per Unit	102.60	25.50	133.80	51.10	
Traditional Cost System (Direct Labor Hours)					
Total Overhead	227.82	2,278.18	683.45	6,834.55	10,024.00
Cost Per Unit	22.78	22.78	68.35	68.35	
Accuracy Per Product	798.18	271.82	654.55	(1,724.55)	
Weighted Products with Cost Driver 2 Only					
Total Overhead	589.65	2,358.59	1,179.29	5,896.47	10,024.00
Cost Per Unit	58.96	23.59	117.93	58.96	
Accuracy Per Product	436.35	191.41	158.71	(786.47)	
Nonweighted Products (Cost Driver 2 with 1 and 3, and 4 with 5)					
Total Overhead	801.41	2,805.65	1,202.82	5,214.12	10,024.00
Cost Per Unit	80.14	28.06	120.28	52.14	
Accuracy Per Product	224.59	(255.65)	135.18	(104.12)	

[a] Overhead consumption intensity is computed with direct labor hours as follows: Intensity = Total costs / Total direct labor hours = 1Q 024 / 220 = $45.56.

[b] The accuracy for any given product and pair of cost drivers is the difference between (1) product cost when all the cost drivers are used and (2) the cost after the combination is applied.

IV. Conclusion

In this study, we analyzed the two interrelated decisions regarding the optimal number and the selection of cost drivers to be included in an activity-based costing system. We started with the realization that information carrying and processing costs are asssociated with costs drivers. Corporate outlays for cost drivers also include the related costs necessary for planning and control; with an excessive number of cost drivers, the costs incurred for planning and control can be prohibitive. Thus, on a cost-benefit theme, a balance must be sought between the accuracy benefits derived from multiple cost drivers and the cost of data collection, storage, and processing associated with these drivers. We select the optimal cost drivers, not only for product costing, but also for planning and control.

We formulated this decision process as a model and provided an efficient algorithm to determine the optimal number of cost drivers that capture most of the incurred costs. The algorithm also identifies a priority order, according to which low-priority and relatively insignificant activities with their associated drivers will be combined to save costs without sacrificing much accuracy. Implications and limitations of the model have been provided, as well as a suitable illustration.

The cost model we selected is quite simple and can be extended in various ways. The interrelationships between cost drivers can be specified to recognize specific savings with particular combinations of cost drivers. For example, if cost drivers k and m are derived from one computerized system, while cost driver n is derived from a different sytem, the savings when combining k and m may be different from that with a combination of k and n. In such a case, c_{km} and c_{kn}, rather than a single c_k are considered. Other elements of cost and savings may be included; in particular, a term that relates to users ability to comprehend multiple cost drivers may be included. In this case, the objective function to be optimized becomes a utility function, rather than dollar savings; it still can be used to determine a desired strategy for the organization. Finally, note that our approach can be applied whenever a balance has to be made between the value of increased accuracy, through the collection and processing of additional data, and the costs associated with the accumulation and processing of these data.

The products considered in this study can be replaced by any other cost objects. For instance, if profitability in customer or distribution channels is of interest, the U_i could be redefined to be associated with the related cost objects. In this way the ABC system is extended to include other needs of activity-based management. Similarly, one can also analyze the costs of poor quality, and understand the monetary effects of quality-related decisions.

References

Berliner, C., and J. A. Brimson. 1988. *Cost Management for Today's Advanced Manufacturing: The CAM-I Conceptual Design*. Boston, MA: 1988, Harvard Business School Press.

Cooper, R. 1988a. The rise of activity-based costing part one: What is an activity-based cost system? *Journal of Cost Management* 2 (2): 45–54.

———. 1988b. The rise of activity-based costing part two: When do I need an activity-based cost system? *Journal of Cost Management* 2 (3): 41–48.

———. 1989a. The rise of activity-based costing part three: How many cost drivers do you need, and how do you select them? *Journal of Cost Management* 2 (4): 34–46.

———. 1989b. The rise of activity-based costing part four: What do activity-based cost systems look like? *Journal of Cost Management* 3 (1): 34–46.

———, and Kaplan, R. S. 1992. Activity-based systems: Measuring the cost of resource usage. *Accounting Horizons* 6 (3): 1–13.

Dewan, R. M., and R. P. Magee. 1992. Product costing systems: Information, economics, and implementation. Working paper, Northwestern University, Evanston, IL.

Feltham, G. A. 1977. Cost aggregation: An information economic analysis. *Journal of Accounting Research* 15 (Spring): 42–70.

Garey, M. R., and D. S. Johnson. 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York: W. H. Freeman.

Kaplan, R. S. 1983. Measuring manufacturing performance: A new Challenge for managerial accounting research. *The Accounting Review* 58 (October): 686–705.

———, 1984. Yesterday's accounting undermines production. Harvard Business Review (July–August): 95–101.

———. 1988. One cost system is not enough. *Harvard Business Review* (January–February): 61–66.

Martello, S., G. Laporte, M. Laporte, M. Minoux, and C. Ribeiro, eds. 1987. *Surveys in Combinatorial Optimization, Annals of Discrete Mathematics* vol. 31. Amsterdam, The Netherlands: North-Holland.

———, and P. Toth. 1988. A new algorithm for the 0–1 knapsack problem. *Management Science* 34 (May): 633–44.

O'Guinn, M. 1990. Focus the factory with activity-based costing. *Management Accounting* vol. LXXII. (February): 36–41.

———. 1991. *A Complete Guide to Activity-Based Costing*. Englewood Cliffs, NJ: Prentice Hall.

Rotch, W. 1990. Activity-based costing in service industries. *Journal of Cost Management* 4 (2): 4–14.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 576–601

An Empirical Study of Cost Drivers in the U.S. Airline Industry

Rajiv D. Banker
University of Minnesota
Holly H. Johnston
Rice University

SYNOPSIS AND INTRODUCTION: Recent research on cost driver analysis by Miller and Vollman (1985) and Cooper and Kaplan (1987) suggests that transactions deriving from the diversity of a firm's product line and the complexity of its production process, in addition to output volume, drive overhead costs. As a consequence, it is argued, conventional cost accounting systems based only on volume-related measures, such as units of output, direct labor hours, or machine hours, produce biased and materially misleading cost estimates for managerial decisions on price and product line (whether to continue or discontinue products, or to offer additional products). Systematic biases in cost estimates may also lead to distortions in flexible budgeting systems, variance analyses, and responsibility-accounting systems. Perhaps more important in the long run, omission of operations-based cost drivers may distort the investigation of the likely effects on costs of changes in operating strategies.

Many firms have moved ahead on the basis of this perceived need for more accurate cost estimates and have designed and implemented activity-based costing systems (Schiff 1991). From an academic perspective, however, there is a need for further formal empirical research in this field. Cooper and Kaplan's (1987) evidence is based on field-study discussions with managers in a variety of manufacturing settings and experimentation with cost allocation and product-costing systems based on transactions.

Foster and Gupta (1990) provide some of the first empirical evidence on the correlation of manufacturing overhead with output volume and operations-based measures that reflect characteristics of the manufacturing process. Using data obtained from 37 plants of a single manufacturing firm, Foster and Gupta found that most of the volume-related measures of output

Helpful comments and suggestions from four anonymous referees, an ad hoc associate editor, and seminar participants at the 1989 American Accounting Association annual meeting, Carnegie Mellon University, Dartmouth College, University of Florida, University of Maryland, Massachusetts Institute of Technology, Michigan State University, University of Minnesota, Naval Postgraduate School, Northeastern University, Pennsylvania State University, Rice University, San Diego State University, University of Southern California, and Stanford University are gratefully acknowledged.

Submitted April 1990.
Accepted October 1992.

were highly correlated with manufacturing overhead (MOH), but because only a few measures of manufacturing complexity and efficiency were highly correlated with MOH, their findings leave the impression that systems based on just volume may not significantly distort information generated for managerial decision making.[1] In contrast, we find empirical evidence in favor of incorporating operations-based cost drivers along with measures of volume in cost driver models.

We draw upon previous work in cost accounting and economics to develop analogs in the airline industry for product diversity, production run volumes, and process complexity, and propose a framework for cost driver analysis in the U.S. airline industry. Using a panel of quarterly data for 1981-1985 compiled primarily from traffic and financial statistics submitted by carriers to the Civil Aeronautics Board (CAB) and Department of Transportation (DOT), we specify and estimate a multivariate system of cost functions with multiple cost drivers for the industry during the transition following deregulation. We find both volume- and operations-based cost drivers to be statistically significant. We also demonstrate the potential managerial importance of the operations-based drivers by explaining variations in marginal costs across airlines in terms of operating strategies reflected in the cost driver values.

Empirical cost driver analysis is managerially significant for the industry and period that we examine. The proportion of indirect costs is large, and identification of input consumption for specific services is difficult. During the transition following deregulation, carriers adopted a rich variety of strategies to improve productivity, reduce costs, and increase market share. These strategies directly involved both volume- and operations-based cost drivers. The analytical framework and model that we have developed on the basis of prior literature concerned with the airline industry enable us to examine the differential cost effects of some of the most important strategies adopted.

Key Words: *Cost drivers, Volume-based drivers, Capacity, Operations-based drivers, Cost estimation, Airline industry, Hub-and-spoke strategy, Benchmarking, Variance analysis.*

Data Availability: *Airline cost and production data can be purchased from the Department of Transportation. Data used for this study are available from the authors to researchers willing to share the cost of acquisition.*

THE article is organized as follows. The cost driver model is developed in section I, and the data and estimation procedure are described in section II. In section III, we discuss the empirical findings, focusing on the statistical significance of the operations-based cost drivers that represent strategic choice variables. Section IV contains a discussion of the cost implications of operating choices and an explanation of the variations in the airlines' marginal costs in terms of the cost drivers. Concluding remarks are offered in section V.

---

[1] Foster and Gupta (1990, 310) state that "there is not a strong association between complexity or efficiency variables and MOH levels across the 37 facilities." Banker et al. (1992), however, show that Foster and Gupta's partial correlation approach will tend to bias the test statistic downward, which makes the rejection of the null hypothesis of no association less likely than indicated by the apparent significance level.

I. Cost Driver Model

We model the airline industry cost function as a linear multivariate system of equations to test the importance of volume- and operations-based cost drivers. The model has ten equations, one for each of ten distinct input cost categories. The dependent variables are physical units of inputs, such as labor hours and gallons of fuel, or input costs deflated by appropriate indices. Each equation has its own set of volume- and operations-based drivers as independent variables, depending on theoretical considerations appropriate to the corresponding input. Cost estimates are obtained by multiplying the physical quantities by appropriate prices.

The system of equations is additive and separable with respect to inputs, as in a Leontief-type production process, to reflect the limited possibilities for substitution between inputs that exist in the production of air transportation services once management chooses production arrangements and technology. Managerial choices include those pertaining to aircraft, network configuration, hub concentration, and flight frequency. Direct substitution between inputs, such as pilot labor and ground equipment, or flight attendant labor and fuel, cannot occur once these choices have been made. Indirect substitution between inputs is possible, however, as managers can effectively vary the input proportions by suitably varying their choices of aircraft and production arrangements.

Cost Categories

The input cost categories listed below are subsets of traditional broad economic categories and are based on the operating expense categories of the Uniform System of Accounts used by carriers to file financial data with the CAB and DOT. The dependent variables are physical measures of the inputs when these are available, and deflated cost measures otherwise.

1. Fuel, gallons of jet fuels and oils.
2. Flying operations labor, hours of labor of flight crews, including pilots, copilots, navigators, and flight engineers.
3. Passenger service labor, hours of labor of flight attendants.
4. Aircraft traffic and servicing labor, hours of labor of ground personnel servicing aircraft and handling passengers at gates, baggage, and cargo.
5. Promotions and sales labor, hours of labor of reservations and sales agents primarily, but also of personnel involved in advertising and publicity.
6. Maintenance labor, hours of labor involved in maintenance of flight equipment and ground property and equipment.
7. Maintenance materials and overhead, total cost of maintenance of property and equipment, deflated by the Producer Price Index for fabricated metals.
8. General overhead, total expenses corresponding to supplies, general and administrative personnel, utilities, insurance, communications, and the like, deflated by the GNP Implicit Price Deflator.
9. Ground property and equipment, flows of service from ground property and equipment, calculated with the method developed by Christensen and Jorgen-

son (1969)[2] and including landing fees deflated by the Air Transport Association cost index for landing fees and rental expenses for ground property and equipment deflated by the Producer Price Index for fixed nonresidential structures.

10. Flight equipment, flows of service from flight equipment (airframes, aircraft engines, avionics, etc.), calculated by imputing fair market rental values deflated by the Producer Price Index for fixed-wing aircraft to owned and leased aircraft by aircraft categories.

Industry means of operating costs by category and descriptive statistics for physical measures of inputs and volume- and operations-based drivers are presented in the appendix.

Volume-Based Cost Drivers

The volume- and operations-based drivers for each input category are listed in table 1. In cost accounting, measures of actual outputs are typically used as cost drivers. The economics literature on airline production has followed this practice by using total revenue passenger and ton miles[3] as the volume-based determinant of total costs (Caves et al. 1984; Kirby 1986; Sickles 1985; Sickles and Good 1986). However, it appears that only the costs for personnel who handle passengers and cargo and for those who handle reservations and sales vary in direct proportion to *actual outputs*, that is, to the number of passengers and tons of cargo actually handled. Instead, most costs appear to vary in proportion to *output capacity*, the amount of seating and cargo capacity available. (Carriers offer seating and cargo capacity, on particular flights and aircraft, for sale in the marketplace, but the actual quantities exchanged are determined by the interaction of the supply of capacity and demand.) Fuel consumption and labor hours for scheduled flight crews and attendants vary more with aircraft size, seating capacity, distance, and other characteristics of flights (and aircraft) than with the actual number of passengers and tons of cargo carried. Expenditures on ground property and equipment, general overhead, and maintenance overhead are associated more with the airline's overall productive capacity than with revenue outputs. Also, the cost of aircraft maintenance varies more with the number of flights, hours flown, and characteristics such as the number of engines than with revenue outputs.

In the airline industry, output capacity is typically measured in terms of available seat miles (ASM) and available ton miles (ATM), and actual outputs are measured in terms of revenue passenger and ton miles, or the number of passengers and tons of cargo. Since the major airlines are primarily passenger carriers, for convenience and ease of interpretation, we convert the available cargo capacity to available seating capacity and combine the two in terms of *capacity seat miles* (CSM) as a measure of output capacity.[4]

---

[2] The "perpetual inventory" method was developed by Christensen and Jorgenson (1969) to measure the service flow of property assets, and was applied to the airline industry by Caves et al. (1984), Sickles (1985), and Sickles and Good (1986).

[3] A revenue passenger (ton) mile is defined as air transportation service for one passenger and his or her baggage (one ton of cargo) for a distance of one mile.

[4] We use the standard industry conversion factor (one ton of cargo carried one mile is equated with ten passengers and their baggage, at approximately 200 pounds each, carried one mile) to convert available ton miles for cargo to available seat miles and sum to get CSM. Cargo generates a small proportion (5.1 percent on average) of the total revenues for the large air carriers, but tends to be highly correlated with passenger miles (correlation coefficient = 0.86). Therefore, we aggregate the two to avoid collinearity problems in estimating our model.

Table 1
Cost Drivers by Input Category

Input Category and Measurement Units	*Cost Drivers (and Hypothesized Signs)*			
	Volume-Based Drivers		*Operations-Based Drivers*	
Fuel in gallons	CSM by aircraft type	(+)	Average stage length	(−)
Flying operations, labor hours	CSM by aircraft type	(+)	Density	(−)
			Hub concentration	(−)
			Hub domination	(−)
Passenger service, labor hours	CSM by aircraft type	(+)	Density	(−)
			Hub concentration	(−)
			Hub domination	(−)
Maintenance, labor hours	CSM by aircraft type	(+)	Density	(−)
			Hub concentration	(−)
			Hub domination	(−)
			Scale	(+)
Maintenance materials and overhead, deflated costs	CSM by aircraft type	(+)	Density	(+)
			Hub concentration	(−)
			Hub domination	(−)
Aircraft and traffic servicing, labor hours	Passengers	(+)	Density	(+)
			Hub concentration	(−)
			Hub domination	(−)
Promotions and sales, labor hours	Passengers	(+)	Density	(+)
			Hub concentration	(−)
			Hub domination	(−)
General overhead, deflated costs	Total CSM	(+)	Density	(+)
			Hub concentration	(−)
			Hub domination	(−)
			Scale	(+)
Group property and equipment, deflated costs	Total CSM	(+)	Density	(−)
			Hub concentration	(+)
			Hub domination	(+)
			Scale	(+)

Note: CSM, capacity seat mile, is defined as the space to carry one passenger and his or her baggage (or one-tenth of one ton of cargo) one mile.

Since differences in aircraft models are hypothesized to be important in determining input requirements for labor for flying operations, passenger service, and maintenance as well as for fuel and maintenance materials and overhead, we have placed various aircraft models into one of eight categories according to fuel consumption and size of flight crew, as shown in table 2.[5] We then compute CSM on the basis of these categories. The relative magnitudes of the coefficient estimates of output capacities by these categories should be consistent with known differences in capacities, fuel efficiency, and labor requirements for flying operations. In contrast, for ground

---

[5] These categories were developed on the basis of technical descriptions of various aircraft models in the industry publication *Jane's All the World's Aircraft*. They account for: (1) differences in size and capacity, determined by the number of engines, aircraft body width, and seating density; (2) fuel efficiency; and (3) changes in the percentages of the industry fleet over time, especially as these reflect the phasing out of older, less fuel-efficient models and phasing in of newer, more fuel-efficient models.

Table 2
Aircraft Categories and Characteristics

Aircraft Models	*Characteristics*	*Mean ASM Per Gallon*	*Flight Crew*
Older Aircraft, Regular-Bodied			
McDonnell Douglas DC-9-10/15	2-engine turbofan, 1st generation	35.9	2
British Aircraft BAC-111	2-engine turbofan		
McDonnell Douglas CD-9-30/40/50	2-engine turbofan, 2nd generation		
Fokker 28	2-engine turbofan		
Dassault Falcon-20	2-engine turbofan		
Boeing-727-100	3-engine turbofan, 1st generation	33.8	3
Boeing-727-200	3-engine turbofan, 2nd generation	42.2	3
Older Large Aircraft			
McDonnell Douglas DC-10-10/30/40	3-engine turbofan, regular body	51.2	3–4
Lockheed L-1011			
Boeing-707		46.1	3–4
Boeing-720	4-engine turbofan, regular body		
McDonnell Douglas DC-8	4-engine turbofan, regular body		
McDonnell Douglas DC-8-71	4-engine turbofan, wide body		
Boeing-747	4-engine turbofan, regular body		
	4-engine turbofan, wide body		
Newer, Fuel-Efficient Aircraft			
Boeing-737-200/300	2-engine turbofan, regular body	43.0	2
McDonnell Douglas Super-80	2-engine turbofan, regular body	54.6	2–3
Boeing-757	2-engine turbofan, wide body		
Boeing-767	2-engine turbofan, wide body	57.8	2–3
Airbus Industrie A300B2/B4			

Note: ASM is available seat miles.

property and equipment and general overhead, input requirements depend very little on differences in the aircraft operated. Therefore, a general measure of overall size, total CSM, is used as the volume-based driver for these inputs. For the labor of ground personnel who service aircraft, handle passengers at ticket counters and gates, and handle baggage and cargo, the number of passengers is employed as the volume-based driver.[6] Flight equipment inputs are determined (as an identity) by the models and numbers of aircraft operated and the corresponding real rental values.

Operations-Based Cost Drivers

We draw on the airline industry literature to develop operations-based cost drivers for the production of air transportation services. The drivers represent choices of alternative technologies, as embodied in choices of aircraft models, route structures, flight frequency or density, and traffic flow control. These are analogous to drivers such as production run volume, product line diversity, configuration of the manufacturing process, and control of the flow of production, which have been suggested as potential cost drivers in manufacturing contexts (Hayes and Clark 1985; Miller and Vollman 1985).

---

[6] The number of passengers is highly correlated (r=0.95) with the number of flights or aircraft handled.

Aircraft Type. A manufacturer may have more than one production site, assembly line, or available production technology, each with a unique set of characteristics that make it more or less appropriate for production under any given set of circumstances. For example, one assembly line may be labor-intensive and another capital-intensive. If the price of labor were to increase, the manufacturer would substitute capital for labor by producing a larger percentage of its output on the capital-intensive line. Another example relates to machine setup costs. One machine may be relatively flexible and inexpensive to set up between production runs, and another more time-consuming and costly. The machine with lower setup costs would be used primarily for low volume production runs, and the machine with higher setup costs for high volume runs.

Analogously, carriers choose among different types of aircraft to provide service on any given set of routes. The choice essentially involves different production technologies with differing input intensities and relative efficiencies. For example, wide-bodied aircraft such as Boeing 747s can serve densely trafficked long-haul routes efficiently, while smaller aircraft such as Boeing 737s can serve less densely trafficked short-haul routes relatively efficiently. Aircraft choices depend on network characteristics, particularly route length and traffic density, and the availability of particular models of aircraft.[7] The output capacity on each type of aircraft, in turn, determines the required hours of pilot, copilot, flight engineer, navigator, and flight attendant labor, required levels of maintenance, and required quantities of fuel per CSM.

Aircraft Size and Average Stage Length. In a manufacturing environment, economies are associated with large production runs as setup costs are relatively invariant with respect to batch sizes. In air transportation services, output capacity increases with both the number of seats made available and the distance traveled. Thus, there are two concepts analogous to batch size: aircraft size and stage length. If the volume of traffic is heavy enough for a carrier to use larger aircraft on a given flight and route, more CSM can be delivered for a given level of flight crew labor and fuel costs. Both Bailey et al. (1985) and Kirby (1986) have documented decreases in total costs with increases in average aircraft size, all else held constant. Average stage length (ASL) is the average length of a carrier's flights in miles. It has consistently been cited as a potential source of economies. As ASL increases, economies are achieved because fuel consumption is considerably greater during take-off and landing than at cruising altitude and speed (Kirby 1986; Strazheim 1969). Caves et al. (1984) and Kirby (1986) have both found decreases in total costs with increases in ASL.

Density. By increasing the number of flights over its network, that is, the density with which the carrier services its network, a carrier can offer a more diversified set of services. The proliferation of flights may come about for two reasons: (1) efforts to provide a more attractive schedule of flights (i.e., a full range of products so customers do not go to other suppliers) and (2) efforts to utilize productive capacity more fully. In the economics literature, the conventional wisdom is that a carrier can utilize inputs more efficiently by operating more flights or carrying more traffic over a given network. There is some empirical evidence, based on a model with a single measure of revenue outputs and highly aggregated categories of inputs, that economies of density obtain (Caves et al. 1984). However, the underlying dynamics in terms of increased production complexity are not addressed.

---

[7] During the transition following deregulation, in response to increasing competition, changing route structures, and the availability of several newly introduced models of more fuel-efficient aircraft, there was a considerable amount of substitution between aircraft. The percentage of ASM on wide-bodied aircraft declined from 35 percent to 25 percent. The percentage provided with new models increased from approximately 8 percent to 27 percent. However, the adjustment was hampered by long lags between orders and deliveries of new aircraft.

In contrast, Cooper and Kaplan (1987) recognize that, in efforts to utilize productive capacity more fully and provide a fuller range of products to meet customer needs, manufacturers incur additional overhead. By scheduling more flights over a given network segment, a carrier incurs additional setup costs for each additional flight, in terms of handling aircraft on the ground and enplaning and deplaning passengers and cargo. At the same time, it may be able to better utilize its fixed ground property and equipment and general overhead inputs. For overall economies to obtain, the gains from utilizing ground capacity more fully must outweigh the costs of additional support activities.

Hub Concentration. Reconfiguring production systems can also generate important cost savings. By concentrating production at sites geared toward high volume or throughput, a manufacturer may add to general overhead costs (in terms of increased need for support department inputs, additional supervisory personnel, etc.) but achieve net cost savings because of more efficient use of other inputs.

Organizing networks as hub-and-spoke systems was one of the most important strategies adopted by the airlines during the transition following deregulation (Bailey et al. 1985; Borenstein 1992; DOT 1990; GAO 1985, 1990; Graham and Kaplan 1982). Most carriers developed hub airports and structured their route systems for the arrival and departure of many flights within a few hours of each other, with passengers and cargo exchanging planes in between. Carriers can thus achieve substantial economies, for example, in maintaining and repairing their fleets, in using ground property, equipment, and labor, and by filling larger aircraft on hub-to-hub routes. However, to achieve these economies, it is likely that carriers also must use more administrative and supervisory labor for communications and other support services.

If a manufacturer cannot effectively control the arrival of raw materials and orders for finished goods, there will be periods of congestion and slack in production. This is particularly true when shared facilities must be used to produce several different products or services, with each primarily under the control of a different department or supervisor. Analogous effects are likely to occur in the airline industry, particularly at major hubs where there is intense competition for the use of air traffic control and shared ground facilities, which often results in overscheduling, congestion, and delays.

The magnitudes of the economies a carrier can obtain by concentrating flights through hubs are likely to depend on whether the carrier has some monopoly power, as reflected in dominant market shares, at its hub airports (DOT 1990; GAO 1985, 1990). During 1981–1985, USAir and Piedmont carried 60 to 80 and 60 to 67 percent, respectively, of industry traffic through their hubs. In contrast, United, American, Delta, and Eastern faced stiff competition, often with each other, at their hub airports. For example, at Dallas-Fort Worth, American and Delta competed with Braniff and other carriers for 29 to 52 and 16 to 29 percent shares of traffic, respectively. USAir and Piedmont were thus in a position to schedule their flights relatively more efficiently than United, American, Delta, and Eastern.

Scale. The scale of production is often suggested as a potential source of economies and, as such, may also be thought of as a potentially important operations-based cost driver. Research reported to date, though, suggests that economies of scale are either very small or do not obtain for the airlines. White (1979), Caves et al. (1984), and Sickles (1985) have found nearly constant returns to scale. In a cost accounting framework of constant marginal costs, increasing returns to scale are indicated when fixed costs are present because average costs decrease with increasing levels of outputs. Therefore, ground property and equipment, general overhead, maintenance labor, and maintenance materials and overhead inputs, which have fixed cost components, are likely to have increasing returns to scale associated with them, but constant returns to scale are likely to obtain for other cost categories.

Summary

Operations-based cost drivers are introduced into the equations after multiplying them by total CSM or the number of passengers to estimate their effects on input consumption per unit CSM or passenger. The hypothesized relations between the operations-based drivers and different input requirements are summarized below.

1. For aircraft type and size, the relative magnitudes of the coefficient estimates for CSM by aircraft category are examined to test hypotheses concerning returns to aircraft size. The wide-bodied and newer, efficient aircraft categories are hypothesized to have lower coefficients.
2. Average stage length is measured as the ratio of airborne miles flown to the number of flights. Included in the equation for fuel costs, it is hypothesized to have a negative coefficient.
3. Density is measured as the number of flights provided relative to the number of airports served. The coefficient of this ratio is hypothesized to be positive (decreasing returns) for aircraft and traffic servicing labor and for promotions and sales labor. For ground property and equipment and general overhead, the coefficient is hypothesized to be negative (increasing returns).
4. To capture returns from concentrating flights through hubs, we introduce the following variables into the equations for aircraft and traffic servicing, promotions and sales, flying operations, passenger service, and maintenance labor, maintenance materials and overhead, ground property and equipment, and general overhead. First, we consider the percentage of the carrier's own flights that are routed through its own hub airports where the carrier has considerable market share, 60 or more percent of the total number of flights by all carriers through the airport during a quarter. This cutoff provided a stable classification of hubs over time and is consistent with characterizations in the airlines literature (Bailey et al. 1985; GAO 1990). Second, we consider the percentage of the carrier's own flights routed through its hub airports where the carrier faces substantial competition (the carrier has less than 60 percent of the flights through the airport). The coefficients are hypothesized to be positive for general overhead and negative for the remaining inputs. The coefficients for the percentages of flights through dominated hubs are hypothesized to be greater in absolute magnitude than those for competitive hubs.

5. To capture scale, intercept terms are explicitly included in the equations for ground property and equipment, general overhead, maintenance labor, and maintenance materials and overhead. Positive intercepts are hypothesized to reflect increasing returns to scale.[8]

System of Equations

The empirical model is described by the following system of equations:

Fuel: $y_{m1t}=(\Sigma_{j=1}^{8}\beta_{1j}r_{mjt}+\gamma_{1,4}z_{m4t})x_{mt}+\epsilon_{m1t},$

Flying Operations Labor: $y_{m2t}=(\Sigma_{j=1}^{8}\beta_{2j}r_{mjt}+\Sigma_{k=1}^{3}\gamma_{2k}z_{mkt})x_{mt}+\epsilon_{m2t},$

Passenger Service Labor: $y_{m3t}=(\Sigma_{j=1}^{8}\beta_{3j}r_{mjt}+\Sigma_{k=1}^{3}\gamma_{3k}z_{mkt})x_{mt}+\epsilon_{m3t},$

Maintenace Labor: $y_{m4t}=\beta_{40}+(\Sigma_{j=1}^{8}\beta_{4j}r_{mjt}+\Sigma_{k=1}^{3}\gamma_{4k}z_{mkt})x_{mt}+\epsilon_{m4t},$

Maintenance Materials and Overhead: $y_{m5t}=\beta_{50}+(\Sigma_{j=1}^{8}\beta_{5j}r_{mjt}+\Sigma_{k=1}^{3}\gamma_{5k}z_{mkt})x_{mt}+\epsilon_{m5t},$

Aircraft and Traffic Servicing Labor: $y_{m6t}=(\beta_{61}+\Sigma_{k=1}^{3}\gamma_{6k}z_{mkt})s_{mt}+\epsilon_{m6t},$

Promotions and Sales Labor: $y_{m7t}=(\beta_{71}+\Sigma_{k=1}^{3}\gamma_{7k}z_{mkt})s_{mt}+\epsilon_{m7t},$

General Overhead: $y_{m8t}=\beta_{80}+(\beta_{81}+\Sigma_{k=1}^{3}\gamma_{8k}z_{mkt})x_{mt}+\epsilon_{m8t},$

Ground Property and Equipment: $y_{m9t}=\beta_{90}+(\beta_{91}+\Sigma_{k=1}^{3}\gamma_{9k}z_{mkt})x_{mt}+\epsilon_{m9t},$

Flight Equipment: $y_{m0t}=\Sigma_{j=1}^{8}v_{mjt}n_{mjt},$

where,

y_{mit} = units of input i used by carrier m during quarter t, $i=1,2,\ldots,10$,
x_{mjt} = capacity seat miles provided on aircraft type j, $j=1,2,\ldots,8$,
$r_{mjt}=x_{mjt}/x_{mt}$ = proportion of total output capacity provided by aircraft type j, where $x_{mt}=\Sigma_{j=1}^{8}x_{mjt}$,
s_{mt} = revenue passengers enplaned during quarter t,
z_{m1t} = measure of concentration of flights through competitive hubs,
z_{m2t} = measure of concentration of flights through dominated hubs,
z_{m3t} = measure of density of traffic over a network,
z_{m4t} = average stage length,
v_{mjt} = rental value of aircraft type j,
n_{mjt} = number of aircraft type j in carrier m's fleet, and
ϵ_{mit} = random error term for equation i and quarter t, $i=1,2,\ldots,9$.

The coefficients are estimated for the industry as a whole rather than by firm since the time series for each carrier is relatively short. To investigate whether the parameters differ between carriers, we also estimated the system of equations separately for small and large firms by splitting the sample equally and by separating out the four largest carriers. Estimation results for the four subsamples are similar to those reported for the entire sample and provide the same insights about the importance and relative magnitudes of the operations-based drivers. In practice, a firm can collect more detailed and frequent data to estimate its own standard input requirements.

---

[8] When we included intercept terms in the equations for the other inputs to test hypotheses that constant returns to scale obtain for these inputs, we found none of the intercepts to be significantly different from zero.

II. Data and Estimation

The model is estimated by using a panel of quarterly data from the first quarter of 1981 through the fourth quarter of 1985 for 28 major, national, and large regional carriers (see table 6 for a list). There are observations for all 20 quarters for 24 of the carriers and fewer for the remaining four. The unit of analysis is the domestic operating system of each carrier.[9]

The data are derived primarily from traffic and financial statistics from Form 41 reports submitted by certificated carriers to the CAB and the DOT. They include annual labor inputs by categories of operating functions and inventories of carriers' aircraft fleets, supplemented by semiannual and quarterly update information from Form 41 reports, fleet data published in *Air Transport World*, and information gathered from carriers' annual reports. The measures of hub concentration, market share, and monopoly power are developed from the Form 41 schedules. The data were cleaned by extensively cross-checking calculations, identifying outliers in trended data for each firm, and correcting errors by checking original hard copies of data obtained from the DOT.

As we carried out the estimation, we examined the data for collinearity and the residuals for evidence of violations of the ordinary least squares (OLS) assumptions that can result in inefficient estimates of the regression coefficients or biased and inconsistent estimates of their variances, including serial and contemporaneous correlation and heteroscedasticity.

Serial correlation is to be expected since the observations are quarterly and the effects of random shocks could be expected to last longer than one quarter. Also, technical inefficiency, to the extent that it obtains, is reflected in the error term and can be expected to persist across quarters. There was strong evidence of serial correlation among the residuals for each airline, with estimates of autocorrelation coefficients ranging from an average of 0.58 for maintenance materials and overhead to 0.89 for fuel. To correct for its effects, first-order autocorrelation coefficients for each carrier were estimated by a variant of the Prais-Winsten estimator proposed by Park and Mitchell (1980). The data were transformed, including the first observations for each time series, in the usual manner, and a second set of regressions was run with the transformed data.[10] The Park-Mitchell estimator is consistent and has been found to perform well for short time series and trended data.[11]

Contemporaneous correlation between the residuals by carrier may be expected because of interrelationships between the operating functions and commonalities within firms. Random shocks affecting one function could have similar or related effects on other functions. There could be similar reactions across functions to events, internal or external to the organization, that have not been modeled (e.g., major changes in management, bitter labor negotiations or strikes, mergers and acquisitions, or bankruptcy proceedings). In a similar vein, there may be some correlation between residuals across carriers.

---

[9] Only the carriers' domestic operating systems were deregulated by the Airline Deregulation Act. International operations were regulated separately, with route and fare agreements negotiated by treaty.

[10] Park and Mitchell (1980) and Doran and Griffiths (1983) found that estimators using all observations provided substantially more efficient estimates of the autocorrelation coefficients than estimators that do not transform the first observation. They found that estimators omitting the first observations with trended data exhibited very low efficiency, often lower than OLS.

[11] It reduces the extent to which the autocorrelation coefficient tends to be underestimated (see Kmenta and Gilbert 1970). All of the estimators tested resulted in underestimated standard errors, which leaves a substantial probability of making a Type I error, although their variants of the Prais-Winsten estimator performed better in this regard than most other estimators. Therefore, we use stringent levels of probability for testing hypotheses and drawing inferences.

If either of these forms of contemporaneous correlation are present, the coefficient estimates of *separate* regressions are unbiased and consistent but inefficient, and the estimates of their variances could be biased (Parks 1967). However, we found that the correlations between residuals were small, and therefore gains in efficiency from estimating the equations as a set of seemingly unrelated regressions (SUR) (Zellner 1962) were small.[12] Estimating the model as SUR resulted in only minor changes in the coefficients and decreases in some of the estimates of the standard errors. The system weighted R^2 was 95.6 percent. Most coefficient estimates were within one standard error of those from the separate regressions, and test results were all very similar to those reported here.

Heteroscedasticity could be expected, as the observations for carriers operating at larger scales could have larger variances. Breusch-Pagan (1979) and Goldfeld-Quandt (1972) tests provided somewhat conflicting evidence regarding the presence of heteroscedasticity, and the precise form of the process generating the disturbances was not clear. Therefore, we used White's (1980) procedure to correct for heteroscedasticity and obtain consistent estimators of the variance-covariance matrices. Analyses of the residuals following each stage of estimation did not reveal any evidence of remaining serial correlation, heteroscedasticity, or nonlinearity.

Finally, collinearity did not appear to be a serious problem after transforming the data for serial correlation. Collinearity will make tests of significance of operations-based drivers conservative and estimates of coefficients unstable. The coefficient estimates and corresponding variances were stable with respect to small perturbations in the data. Belsley et al. (1980) condition indices were less than 30 for all equations except passenger service labor, maintenance labor, and maintenance materials and overhead. For these three equations, the proportions of the coefficient variances associated with the characteristic roots were all less than 0.5 except for CSM on Boeing 727-200s and the percentage of flights through competitive hubs.

III. Empirical Results and Discussion

Summaries of the results of the regressions on the transformed data are presented in table 3. The model containing both volume- and operations-based drivers appears to fit very well and provides a good representation of the industry production correspondence. With the exception of ground property and equipment, the percentages of varia-

---

[12] There are no efficiency gains from estimating SUR for equations in which all of the independent variables are the same, and the efficienty gains from estimating SUR increase as contemporaneous correlation increases (see Judge et al. 1985; Kmenta 1971). Doran and Griffiths (1983) find lower gains in efficiency when data are trended and there is greater correlation among the explanatory variables across equations. We have several sets of equations with similar sets of explanatory variables. Pearson correlation coefficients among the explanatory variables, by quarter, often are greater than 0.50, and many range from 0.80 to 0.97. The correlations among the residuals across equations, however, tend to be much smaller, with only five of 36 being greater than 0.45. Potential gains in efficiency are also diminished by the small number of quarters for which we have data and the fact that many of our explanatory variables have trended data.

Table 3
Regression Results
(Coefficient Estimates and *t*-Statistics)

Panel A. Equations Containing CSM by Aircraft Category:

		Labor			
	Fuel	*Flying Operations*	*Passenger Service*	*Maintenance*	*Maintenance Materials and Overhead*
Intercept (β_{i0})	—	—		2407	2.30×10^6
	—	—		(0.06)	(10.95)***
C9 (β_{i1})	22.21	0.2299	0.2642	0.3783	2.4396
	(21.06)***	(16.05)***	(14.36)***	(4.73)***	(6.93)***
B727–100 (β_{i2})	28.82	0.2511	0.2619	0.4637	3.1131
	(13.06)***	(5.80)***	(4.05)***	(2.84)**	(2.84)**
B727-200 (β_{i3})	21.18	0.2382	0.3244	0.3291	1.5895
	(33.68)***	(9.12)***	(6.54)***	(3.87)***	(2.49)*
DC-10, L1011 (β_{i4})	17.46	0.1425	0.2243	0.2136	1.8999
	(13.12)***	(5.95)***	(5.84)***	(3.86)***	(3.69)**
B747 (β_{i5})	16.73	0.1056	0.1736	0.2572	1.3866
	(14.92)***	(4.53)***	(5.11)***	(5.26)***	(3.00)**
B737 (β_{i6})	19.30	0.1324	0.2719	0.1929	0.2856
	(48.21)***	(4.10)***	(10.20)***	(1.91)	(0.58)
B757, MD-80 (β_{i7})	16.06	0.0958	0.1321	0.2318	3.9511
	(19.58)***	(2.38)*	(2.08)*	(1.80)	(3.04)**
B767, A300 (β_{i8})	16.55	0.1696	0.3404	0.2507	0.9026
	(14.47)***	(5.24)***	(5.55)***	(2.24)*	(0.38)
Competitive Hub by KCSM (γ_{i1})	—	−0.00108	−0.00077	−0.00162	−0.00977
	—	(−3.42)**	(−1.44)	(−1.92)	(−1.13)
Dominant Hub by KCSM (γ_{i2})	—	−0.00123	−0.00119	−0.00355	0.00297
	—	(−3.49)**	(−2.81)**	(−3.71)**	(0.30)
Density by KCSM (γ_{i3})	—	−0.00001	−0.00002	−0.00003	0.00005
	—	(−1.73)	(−2.01)*	(−0.94)	(0.27)
ASL by KCSM (γ_{i4})	−0.0042	—	—	—	—
	(−5.02)***	—	—	—	—
R^2 (Full Model)	99.1%	96.7%	96.9%	93.2%	94.8%
Pr>F (Full Model)	0.0000	0.0000	0.0000	0.0000	0.0000
R^2 (Volume-Based Drivers Only)	98.9%	92.7%	96.3%	88.0%	94.5%
Pr>F (Operations-Based Drivers)	0.0000	0.0000	0.0133	0.0014	0.1172

Note: Traffic by aircraft type measured in KCSM.

KCSM: Thousand capacity seat miles.
ASL: Average stage length

* Significant at 5 percent level
** Significant at 1 percent level
*** Significant at 0.01 percent level

Pr>F (Full Model): Pr>F ($\beta_{ij}=0$, *for* $j=1,\ldots 8$, $\gamma_{ik}=0$ *for* $k=1,\ldots 4$)
Pr>F (Operations-Based Drivers): Pr>F ($\gamma_{i1}=\gamma_{i2}=\gamma_{i3}=\gamma_{i4}=0$)

Table 3—*Continued*

Panel B. Equations Containing Total CSM and Passengers:

	Labor			
	Aircraft and Traffic Service	*Promotions and Sales*	*General Overhead*	*Ground Property and Equipment*
Intercept (β_{i0})	— —	— —	20×10^6 (17.28)***	76×10^6 (15.53)***
Total KCSM (β_{i1})	— —	— —	4.583 (6.06)***	15.269 (2.04)*
Total KP (β_{i1})	905.521 (7.89)***	402.973 (7.71)**	— —	— —
Competitive Hub by KCSM (γ_{i1})	— —	— —	−0.049 (−2.67)*	−0.226 (−2.43)*
Competitive Hub by KP (γ_{i1})	−9.013 (−3.51)**	−3.686 (−2.86)**	— —	— —
Dominant Hub by KCSM (γ_{i2})	— —	— —	−0.058 (−2.92)**	−0.260 (−2.13)*
Dominant Hub by KP (γ_{i2})	−14.795 (−5.29)***	−8.335 (−6.23)***	— —	— —
Density by KCSM (γ_{i3})	— —	— —	−0.0001 (−0.20)	−0.0008 (−0.32)
Density by KP (γ_{i3})	0.066 (1.87)	0.050 (2.60**)	— —	— —
R^2 (Full Model)	87.3%	88.8%	96.9%	24.0%
Pr > F (Full Model)	0.0000	0.0000	0.0000	0.0000
R^2 (Volume-Based Drivers Only)	85.0%	86.6%	96.8%	21.0%
Pr > F (Operations-Based Drivers)	0.0000	0.0000	0.0008	0.0003

Note: KCSM: Thousand capacity seat miles.
KP: Thousand passengers.

* Significant at the 5 percent level.
** Significant at the 1 percent level.
***Significant at the 0.01 percent level.

Pr > F (Full Model): Pr > F ($\beta_{i1}=\gamma_{i1}=\gamma_{i2}=\gamma_{i3}=0$).
Pr > F (Operations-Based Drivers): Pr > F ($\gamma_{i1}=\gamma_{i2}=\gamma_{i3}=0$).

tion in inputs explained by the volume- and operations-based cost drivers range from 88 percent for aircraft and traffic servicing labor to 98 percent for fuel.[13]

The measures representing output volumes are consistently found to be important cost drivers. For all cost categories, the estimates of the coefficients of the volume-based drivers are positive and generally have *t*-statistics that are highly significant. Also, the relative magnitudes of the coefficient estimates for CSM by aircraft categories are generally consistent with the relative capacities and other characteristics of the aircraft.

---

[13] We believe the low explanatory power for ground property and equipment is attributable to inaccuracies associated with calculating a large component of the measure of the dependent variable with the perpetual inventory method, rather than to omitted variables or measurement errors in the explanatory variables.

The estimated coefficients of the operations-based drivers are generally in the expected directions and have *t*-statistics that are significant at fairly high levels. F-statistics comparing the error sums of squares restricting and not restricting the coefficients of all four of the operations-based drivers to zero are significant at levels ranging from 0.0000 to 0.0133 for all but one input category. Therefore, a model based only on volume-based drivers is misspecified and likely to have omitted-variable bias. If the omitted variables are correlated with the included variables, variation in input requirements that should be attributed to the excluded operations-based drivers will be attributed to the included volume-based variables and the estimated coefficients for volume will be statistically biased and inconsistent. Even if the omitted and included variables are not correlated, both the estimates of intercept terms and estimates of the variances of all of the coefficient estimates will be statistically biased and inconsistent (Kmenta 1971; Theil 1957).

As may be seen in table 3, panel A, the coefficient estimate for ASL is significantly negative for the fuel cost category. This lends support to the hypothesis that the marginal requirements for fuel inputs diminish as ASL increases.

For promotions and sales labor, the coefficient for density is significantly positive, which indicates that adding flights on a given network requires additional support labor. The coefficient for passenger service labor is significantly negative. Both results are consistent with the findings of Cooper and Kaplan (1987) about product diversity and illustrate the countervailing forces at work with diversification. The coefficients for flying operations labor, maintenance labor, and maintenance materials and overhead are insignificant. We surmise this is related to standardization due to Federal Aviation Administration (FAA) regulations. The coefficients for ground property and equipment and general overhead are also insignificant and may reflect the inherent difficulties in measuring capital inputs.

The results for competitive and dominated hubs follow similar patterns with large, significantly negative coefficients for labor handling passengers, cargo, and aircraft on the ground. The coefficients for ground property and equipment and for flying operations labor are also significantly negative, and those for general overhead are significantly positive.[14] Thus, by adopting a hub-and-spoke strategy, a carrier can achieve fairly substantial economies in the use of most inputs, but reconfiguring production to a more centralized operation requires increases in general overhead inputs.

The results lend support to the hypothesis that carriers that dominate their hubs, and therefore may have some monopoly power, can achieve relatively greater economies from hub concentration than carriers with competitive hubs. The results of asymptotic χ^2-tests indicate that the coefficients for aircraft and traffic servicing, promotions and sales, and maintenance labor are significantly more negative for dominated than for competitive hubs ($pr > \chi^2 = 0.0378$, 0.0005, and 0.0444, respectively). This suggests that carriers that can dominate hubs may be able to schedule, market, and operate flights in a manner that enables them to control the flow of traffic and thereby use their resources more efficiently.

---

[14] This provides some of the first detailed empirical evidence for the industry as a whole concerning the magnitude of economies from the adoption of hub-and-spoke systems.

Table 4
Differences in Input Requirements Between Aircraft Types
(Asymptotic χ^2-statistics)

Older Wide-Bodied Models	Older Regular-Bodied Models			Newer Fuel-Efficient Models		
	DC-9 BAC-111	B727–100	B727–200	B737	B757 MD-80	B767 A300
Fuel						
DC-10, L-1011	4.7451	11.3508	3.7189	1.8310	−1.4058	−0.9181
χ^2	20.407	89.868	16.966	2.268	1.380	0.129
$Pr>\chi^2$	0.0000	0.0000	0.0000	0.1321	0.2401	0.7197
DC-8, B707, B747	5.4805	12.0862	4.4543	2.5664	−0.6704	−0.1827
χ^2	44.598	368.669	5.914	5.914	0.344	0.007
$Pr>\chi^2$	0.0000	0.0000	0.0150	0.0150	0.5577	0.9340
Flying Operations Labor						
DC-10, L1011	0.087386	0.108611	0.095651	−0.010143	−0.046664	0.027088
χ^2	6.444	12.119	16.188	0.239	3.138	0.267
$Pr>\chi^2$	0.0111	0.0005	0.0001	0.6249	0.0765	0.6056
DC-8, B707, B747	0.124297	0.145522	0.132562	0.026768	−0.009753	0.063999
χ^2	9.437	30.602	30.262	3.409	0.090	2.167
$Pr>\chi^2$	0.0021	0.0000	0.0000	0.0648	0.7641	0.1410

Note: The first entry in each cell is the difference between coefficients (column less row category).

The estimates for the intercepts in the equations for ground property and equipment, general overhead, and maintenance materials and overhead are positive and significant, and therefore lend support to hypotheses that increasing returns to scale obtain for these inputs. The intercept for maintenance labor is not significant. We surmise that, with FAA maintenance requirements, maintenance labor hours become nearly proportionate to output capacity.

The results of asymptotic χ^2-tests for differences in input requirements for fuel and flight crew labor between older large aircraft, older small aircraft, and newer fuel-efficient aircraft are presented in table 4. Comparisons between older large and small aircraft indicate highly significant differences in coefficients and are consistent with the hypothesis and industry evidence that wide-bodied aircraft require less fuel and flying operations labor inputs than older regular-bodied aircraft. However, sharp differences between older large aircraft and newer models are not indicated. This suggests that the improvements in fuel efficiency and reductions in required crew size for the newer models have made up for the size-based advantages larger aircraft had in the past.

IV. Cost Effects of Managerial Strategies

The empirical results in the preceding section demonstrate that the effects of operations-based cost drivers are statistically significant. We next examine the magnitude of these effects from a managerial perspective. Our model is explicitly constructed so that we can estimate the differential effects of changes in the operations-based drivers on marginal costs by cost category. We also show how operations-based drivers can be incorporated into variance analysis relative to industry benchmarks and used for performance evaluation.

Effects on Marginal Costs

The estimated marginal costs (costs per additional 1,000 CSM) for each aircraft category are presented in table 5. The marginal cost for each input i is obtained by multiplying the right-hand side of equation i by the price of input i and taking the first derivative with respect to CSM. For this illustration, we use the sample averages for the input prices, p_i. When different measures of volume are used as primary cost drivers (numbers of passengers as opposed to CSM), conversion factors are used to obtain the required common denominator. Therefore, for aircraft and traffic servicing labor and promotions and sales labor, we multiply the number of passengers by the CSM per passenger ratio to obtain CSM. For flight equipment, the number of aircraft is multiplied by the CSM per aircraft ratio. The columns and subtotals in panel A contain estimates of the components of marginal costs associated with changes in volume for each aircraft type. Except for flight equipment, these estimates do not differ across carriers. Panel B contains estimates of the components of marginal costs that reflect adjustments, for the operations-based drivers for a representative firm, with operations-based drivers taking on the sample means. The estimates do not differ across aircraft categories. The marginal cost totals combine the incremental costs associated with both direct volume-based drivers and corresponding adjustments associated with managerial decisions to alter the operations-based drivers. With the exception of density, the estimated marginal savings (aggregated across inputs) associated with operations-based drivers are significantly greater than zero.[15]

The estimates presented in table 5 highlight the relative magnitudes of the various components of marginal costs, both across aircraft categories and between the volume- and operations-based drivers, and can provide the basis for longer term decisions regarding network structure and fleet mix. The advantages of newer and wide-bodied aircraft in the use of flying operations labor, fuel, maintenance materials and overhead, and maintenance labor are evident (see panel A). When combined, they result in marginal costs ranging from $46 to $51 per 1,000 CSM, as opposed to $63 to $75 for the older regular-bodied models for the representative firm.

The separate estimates for different aircraft categories in panel A, before adjustments for operations-based drivers, are aggregated into a single number by weighting them with the average CSM for their respective categories. The weighted averages of these marginal cost components are presented in the first column of panel B. For the representative carrier, the weighted average incremental cost is approximately $83 per 1,000 CSM (before adjustment for operations-based drivers), but the total operations-based savings are approximately $27, which leaves a net marginal cost of $56. The savings derive largely from concentrating flights through hubs. As may be seen from the descriptive statistics calculated by carrier and quarter (see the lower part of panel B), the potential savings from concentrating flights through hubs average approximately $27 and range from $3 to $139. For carriers dominating hubs, estimated savings range up to $72. The savings associated with increasing ASL are smaller, averaging $2 per 1,000 CSM and ranging up to $8. The savings associated with increasing density are the

---

[15] Standard error estimates, under the assumption of no contemporaneous correlation across inputs, were calculated as the square root of the sum of the estimated variances weighted by the squares of their prices.

Table 5
Marginal Costs By Aircraft Category
(Per 1,000 CSM)

Panel A. Volume-Based Components:

		Older Regular-Bodied Aircraft			*Older Wide-Bodied Aircraft*		*New Fuel Efficient Aircraft*		
Input Category y_i	*Input Price* p_i	*DC-9* $p_i\hat{\beta}_{i1}$	*B727–100* $p_i\hat{\beta}_{i2}$	*B727–200* $p_i\hat{\beta}_{i3}$	*DC-10* *L1011* $p_i\hat{\beta}_{i4}$	*DC-8* *B707, B747* $p_i\hat{\beta}_{i5}$	*B737* $p_i\hat{\beta}_{i6}$	*B757* *MD-80* $p_i\hat{\beta}_{i7}$	*B767* *A300* $p_i\hat{\beta}_{i8}$
Fuel	\$ 1.0041	\$22.3000^a	\$28.9330^a	\$21.2697^a	\$17.5356^a	\$16.7972^a	\$19.3740^a	\$16.1240^a	\$16.6137^a
Flying Operations Labor	41.1801	9.4669^a	10.3410^a	9.8073^a	5.8683^a	4.3483^a	5.4506^a	3.9467^a	6.9838^a
Passenger Service Labor	14.7975	22.3000^a	3.8750^a	4.7999^a	3.3189^a	2.5687^a	4.0227^a	1.9542^a	5.0371^a
Maintenance Labor	31.6610	11.4965^a	14.0930^a	10.0002^a	6.4908^a	7.8170^a	5.8616	7.0450	7.6197
Maintenance Materials & Overhead	1.0000	2.4396^a	3.1131^a	1.5895^a	1.8899^a	1.3866^a	0.2856	3.9511^a	0.9026
Aircraft & Traffic Servicing Labor	18.1737	9.6037^a	9.6037^a	9.6037^a	9.6037^a	9.6037^a	9.6037^a	9.6037^a	9.6037^a
Promotions and Sales Labor	39.1818	9.2141^a	9.2141^a	9.2141^a	9.2141^a	9.2141^a	9.2141^a	9.2141^a	9.2141^a
General Overhead	1.0000	4.5830^a	4.5830^a	4.5830^a	4.5830^a	4.5830^a	4.5830^a	4.5830^a	4.5830^a
Ground Property & Equipment	1.0000	15.2655^a	15.2655^a	15.2655^a	15.2655^a	15.2655^a	15.2655^a	15.2655^a	15.2655^a
Flight Equipment		2.2407	1.9490	3.2615	0.9171	0.8084	3.8498	2.4254	0.2410
Volume-Based Driver Subtotal		\$90.5197	\$100.9700	\$89.3944	\$74.6869	\$72.3925	\$77.5106	\$74.1127	\$76.0642
Standard Error of Subtotal		2.7546	5.8959	3.0276	2.4853	2.2049	3.4208	4.6186	3.9922
Operations-Based Driver Subtotal		−26.5084	−26.5084	−26.5084	−26.5084	−26.5084	−26.5084	−26.5084	−26.5084
Volume- and Operations-Based Total		\$64.2691	\$74.7194	\$63.1438	\$48.4368	\$46.1419	\$51.2600	\$47.8621	\$49.8136
Conversion Factors:									
Mean CSM		537,268	267,443	1,561,244	1,092,739	589,573	382,142	170,596	237,153
Passenger to CSM Ratio	0.00059								
Aircraft to CSM Ratio		0.000036	0.000030	0.000018	0.000009	0.000015	0.000029	0.000017	0.000010

Continued

Table 5—*Continued*

Panel B. Operations-Based Components:

Input Category y_i	Weighted Average Volume-Based Driver Total[b]	Hub Concentration: Competitive $p_i\hat{\gamma}_{i1}z_{R1}$	Hub Concentration: Dominated $p_i\hat{\gamma}_{i2}z_{R2}$	Density $p_i\hat{\gamma}_{i3}z_{R3}$	Average Stage Length $p_i\hat{\gamma}_{i4}z_{R4}$	Operations-Based Driver Total $p_i\Sigma_k\hat{\gamma}_{ik}z_{Rk}$	Volume- & Operations-Based Driver Total
Calculations for Representative Firm:							
Fuel	$19.8599				$−2.4112[a]	$−2.4112	$17.4487
Flying Operations Labor	7.5550	$−1.8572[a]	$ −0.2743[a]	$−0.3576		−2.4891	5.0658
Passenger Service Labor	3.8934	−0.4749	−0.0951[a]	−0.2841[a]		−0.8541	3.0393
Maintenance Labor	8.7862	−2.0519	−0.5837[a]	−0.6747		−3.3104	0.4758
Maintenance Materials & Overhead	1.7579	−0.4064	0.0161	0.0411		−0.3492	1.4087
Aircraft & Traffic Servicing Labor	9.6037	−3.9745[a]	−0.8488[a]	0.5441		−4.2792	5.3245
Promotions and Sales Labor	9.2141	−3.5045[a]	−1.0310[a]	0.8870[a]		−3.6485	5.5655
General Overhead	4.5830	2.0377[a]	0.3155[a]	−0.0856		2.2676	6.8506
Ground Property & Equipment	15.2655	−9.3856[a]	−1.4054[a]	−0.6432		−11.4343	3.8312
Flight Equipment	2.1161						2.1161
Operations-Based Driver Subtotal	$82.6347	$−19.6173	$−3.9068	$−0.5730	$−2.4112	$−26.5084	$56.1264
Standard Error of Subtotal		$4.4776	$0.7326	$2.2533	0.4800		
Mean Driver Level		41.58	4.41	777.80	557.35		
Passenger to CSM Ratio	0.0006						
Descriptive Statistics for Calculations for Sample Firms:							
Mean Savings		$−26.5712	$−4.6922	$−0.9235	$−2.4112	$−33.2377	
Standard Deviation		27.2099	11.4729	3.7597	1.3004	22.7887	
Minimum		−139.0104	−71.9077	−1.5021	−8.3219	−131.7645	
Maximum		−3.2457	0.0000	20.6472	−0.5104	−9.8156	

---

[a] Estimates are based on coefficient estimates significant at 0.05 or higher.

[b] Weights corresponding to CSM by aircraft type are used to average original costs for each input in panel A.

least, averaging \$1 and ranging up to \$1.50. Approximately 38 percent of the observations, however, exhibit net increases in costs with increases in density. The estimated net savings associated with all operations-based drivers average \$33 and range from \$10 to \$132 per 1,000 CSM.

Variance Analysis

The understanding of the cost structure gained in estimating the underlying production function can be useful for planning and in establishing standards for budgeting and variance analyis. Also, *ex post* analysis of variances from industry benchmarks can be useful for performance evaluation and management control. It can be used to evaluate operating strategy choices or, alternatively, to analyze a firm's performance relative to others in its industry. We illustrate such variance analysis between industry benchmarks and actual levels for operations-based drivers, and input prices in terms of their effects on marginal costs.

Variances in marginal costs for firm *m* and period *t* (summed over inputs *i*) may be calculated as:

$$\Delta_{mt} = \sum_i (MC_{Rit} - MC_{mit}),$$

where,

$$MC_{mit} = p_{mit}\left(\hat{\beta}_{i*} + \sum_k \hat{\gamma}_{ik} z_{mkt}\right),$$

p_{mit} is the price of input *i* paid by firm *m* in period *t*, the subscript *R* in place of *m* denotes industry benchmark values, and $\hat{\beta}_{i*}$ is the average of $\hat{\beta}_{ij}$ weighted by CSM for aircraft type *j*. Industry benchmarks are based on weighted averages for the entire sample. The variance is favorable if it is positive, that is, if the marginal cost for firm *m* is less than the benchmark marginal cost. Δ_{mt} may be factored into input price and operations-based driver variances as follows:

$$\Delta_{mt} = \sum_i (p_{Rit} - p_{mit})\left(\hat{\beta}_{i*} + \sum_k \hat{\gamma}_{ik} z_{mkt}\right) \qquad \text{(input price variance)}$$

$$+ \sum_i \sum_k p_{Rit} \hat{\gamma}_{ik} (z_{Rkt} - z_{mkt}) \qquad \text{(operations-based driver variance).}$$

The input price variance captures differences associated with deviations between actual and benchmark input prices. It is favorable if the prices firm *m* pays for its inputs are collectively less than the benchmark prices. The operations-based driver variance captures differences in marginal costs associated with deviations between firm and benchmark driver levels. Since the coefficients of these drivers are negative, the variance is favorable if the firm's operations-based drivers exceed the benchmark levels.

Table 6 contains variances for the sample firms. We used mean rather than quarterly data for each firm and computed the variances relative to values for the representative firm that reflect industry means for the period 1981–1985. The results show considerable variation in marginal costs that can be attributed to operations-based drivers. The standard deviation of \$7.40 in the total operations-based driver variance is large relative to the marginal cost of \$56.13 for the representative firm.

Table 6
Variance Analysis of Marginal Costs of 1,000 Weighted Average CSM
(Firm vs. Industry Means)

	Hub Concentration							
Carrier	*Competitive*	*Dominant*	*Total*	*Density*	*Average Stage Length*	*Operations Driver Subtotal*	*Input Price Variance*	*Total Variance*
AirCal	−0.3750	−2.8190	−3.1939	−0.13583	−0.85083	−4.1806	3.3525	−0.8280
Alaskan	0.7868	−1.4970	−0.7103	−0.30198	−0.07192	−1.0842	−13.2413	−14.3255
Aloha	18.7484	−3.9068	14.8416	0.65010	−1.85727	13.6344	1.1112	14.7456
American	1.5258	−3.9068	−2.3810	0.16468	1.28473	−0.9316	0.7381	−0.1935
Arrow	13.2359	−3.9068	9.3291	−0.55363	2.69359	11.4691	1.6064	10.2330
Braniff	4.7649	−3.9068	0.8580	−0.31494	0.79698	1.3401	8.9135	10.2536
Continental	3.2595	−3.9068	−0.6473	−0.13138	0.78361	0.0049	−0.6579	−0.6530
Delta	−3.2750	−3.9068	−7.1818	0.54160	−0.05437	−6.6946	1.4448	−5.2498
Eastern	−4.2371	−3.9068	−8.1439	0.40355	0.10730	−7.6331	3.9579	−3.6752
Frontier	−3.1833	−3.9068	−7.0901	−0.24935	−0.60747	−7.9469	2.8946	−5.0523
Hawaiian	16.3801	−3.4225	12.9576	0.08595	−1.75845	11.2851	−1.5026	9.7825
Midway	−13.0770	29.3323	16.2553	−0.26101	−0.24728	15.7470	2.8216	18.5686
New York Air	12.2313	−3.9068	8.3245	−0.09199	−1.01429	7.2182	−0.0992	7.1191
Northwest	−2.8749	−3.9068	−6.7818	−0.07769	0.31394	−6.5455	−1.4837	−8.0293
Ozark	0.1396	−3.5397	−3.4001	−0.36971	−0.72970	−4.4995	1.6453	−2.8542
Pacific Southwest	5.1677	−3.9068	1.2608	0.21838	−0.86782	0.6114	−1.8211	−1.4236
PanAmerican	0.8392	−3.9068	−3.0676	−0.31997	1.26505	−2.1226	−19.9864	−22.1089
People Express	8.6514	−3.9068	4.7446	0.07883	−0.24050	4.5829	5.9929	10.5759
Piedmont	−12.9911	6.9010	−6.0901	0.02741	−0.98419	−7.0469	3.3921	−3.6548
Republic	−9.3656	−2.3832	−11.7488	−0.01289	−0.78050	−12.5422	1.4727	−11.0695
Southwest	−12.9847	21.3745	8.3909	0.68047	−1.08479	7.9866	1.3244	9.3110
Texas Int'l	1.4422	−3.9068	−2.4647	−0.36613	−0.47389	−3.3047	5.6353	2.3306
Transamerican	2.2960	−3.2672	−0.9712	−0.42241	1.66495	0.2714	−3.2717	−3.0004
Trans World	−1.2415	−3.9068	−5.1483	0.00797	0.90958	−4.2307	−0.6318	−4.8626
United	−1.9976	−3.9068	−5.9044	−0.01876	1.16767	−4.7555	−4.0696	−8.8251
USAir	−16.1562	15.6956	−0.4606	0.17679	−0.84285	−1.1266	−1.8272	−2.9538
Western	−6.0660	4.5546	−1.5115	−0.07977	0.27156	−1.3197	0.4407	−0.8790
World	11.1592	−3.8908	7.2684	−0.45145	4.88726	11.7042	7.5784	19.2283
Descriptive Statistics for Sample Firms:								
Mean	0.4573	−0.1953	0.2619	−0.0401	0.1314	0.3532	0.2046	0.4468
S. D.	8.9357	8.4789	7.3354	0.3253	1.4151	7.4041	5.7265	9.7732
Minimum	−16.1562	−3.9068	−11.7488	−0.5536	−1.8573	−12.5422	−19.9864	−22.1089
Maximum	18.7484	29.3323	16.2553	0.6805	4.8873	15.7470	8.9135	19.2283

USAir and Republic have similar ASL, USAir has a favorable density variance, and Republic has a favorable input price variance. However, the advantage USAir has over Republic because of its higher percentage of flights through dominated hubs is the most important component of a comparison of the two carriers. Because of this advantage, USAir's marginal costs are estimated to be $8.11 less than those of Republic.

United, the largest carrier during the period, has unfavorable hub concentration and density variances that are partially offset by a favorable ASL variance but leave a net unfavorable operations-based driver variance. United also has an unfavorable input price variance. American, its closest competitor, has the same disadvantages in density and ASL, but is better in hub concentration and input prices, which results in a smaller unfavorable total variance. In contrast, the next two largest carriers, Delta and Eastern, have greater disadvantages in concentration through competitive hubs, which results in more unfavorable operations-based driver variances. Delta and Eastern, however, enjoy favorable input price variances, which make their overall variances less unfavorable.

V. Concluding Remarks

Our findings have demonstrated empirically that, while output capacity and volume are important cost drivers, operations-based drivers related to product diversity and production process complexity are also significant in a major service industry. These drivers often directly reflect management strategies to improve productivity and reduce costs or even to increase market share. Our model, with multiple cost categories and drivers, provides a framework for evaluating the cost effects of these strategies, as well as a method for obtaining cost estimates for other management accounting applications. We have demonstrated empirically that several strategies adopted by U.S. airlines during the transition following deregulation had material effects on costs. Specifically, we found that operations-based drivers reflecting strategies to increase batch size (aircraft size and average stage length) and product diversity (flight density), to reconfigure the production process (increase hub concentration), and to control process flow (develop dominated hubs) have material and differential effects on costs.

Academic research in cost accounting has only recently been directed toward identifying important cost drivers. Although Foster and Gupta's (1990) work with one electronics firm suggests that operations-based drivers are generally not strongly correlated with manufacturing overhead costs, our work provides contrary evidence for an industry. As we begin to accumulate statistical evidence of this type, we will be in a better position to decide whether the traditionally used volume-based drivers need to be augmented with strategy-related operations-based cost drivers. This will, in turn, help management obtain more accurate evaluations of the cost effects of operating decisions.

Table A1

Industry Operating Costs and Physical Measures by Input Category

	Input Costs			*Physical Measures*	
Cost Category	*Mean Level*[a]	*Mean Percentage*	*Standard Deviation*	*Mean Level*[b]	*Standard Deviation*
Aircraft and Traffic Servicing Labor	$ 30,311	8.86%	3.24%	1,724,946	1,980,910
Promotions and Sales Labor	30,285	9.03	2.52	813,849	1,068,803
Flying Operations Labor	30,532	8.57	2.78	664,370	746,424
Passenger Service Labor	16,401	4.60	1.22	1,023,253	1,221,969
Maintenance Labor	19,872	6.95	2.82	933,833	1,309,899
Maintenance Materials and Overhead	7,276	2.06	0.93	7,295,603	9,064,265
General Overhead	37,498	14.62	4.97	37,497,917	39,702,527

Continued

Table A1—*Continued*

Cost Category	Input Costs Mean Level[a]	Input Costs Mean Percentage	Input Costs Standard Deviation	Physical Measures Mean Level[b]	Physical Measures Standard Deviation
Ground Property and Equipment	48,007	12.52	6.38	48,115,431	68,906,924
Flight Equipment	21,269	8.42	3.36	21,268,543	21,581,727
Fuel	75,064	24.37	4.39	85,705,214	95,828,418
Total Costs	$317,075	100.00%	—		

[a] In thousands of 1982 dollars.

[b] Labor in hours, fuel in gallons, and other variables as deflated costs.

Table A2

Industry Volume- and Operations-Based Cost Drivers

Variable	*Mean*	*Standard Deviation*
Volume-Based Cost Drivers (in 1,000s):		
Passengers Enplaned	2,823	2,780
Total CSM	4,838,158	5,792,900
CSM (by aircraft categories):		
B737	382,142	655,774
DC-9	537,268	1,009,699
B757, MD-80	170,596	423,401
B767, A300	237,153	698,257
B727-100	267,443	519,333
B727-200	1,561,244	2,119,059
DC-10, L1011	1,092,739	1,799,622
B747	589,573	1,335,023
Operations-Based Cost Drivers:		
Concentration:		
Competitive Hubs	41.58	19.70
Dominated Hubs	5.41	12.56
Density	777.80	453.06
Average Stage Length	557.35	300.59

Note: CSM = capacity seat miles.

Table A3
Input Costs by Category
(Means in Thousands of 1982 Dollars)

Carrier	*Aircraft & Traffic Servicing Labor*	*Promotions & Sales Labor*	*Flying Operations Labor*	*Passenger Service Labor*	*Maintenance Labor*	*Fuel*	*Ground Property & Equipment*	*Aircraft Rental*	*Maintenance Materials & Overhead*	*General Overhead*
AirCal	$ 6,057	$ 5,992	$ 5,204	$ 2,713	$ 3,621	$ 14,838	$ 5,860	$ 7,368	$ 1,634	$ 9,773
Alaska	7,495	6,362	5,947	2,992	3,879	17,498	10,329	4,921	838	9,901
Aloha	2,477	2,737	1,651	729	2,006	4,330	1,866	3,436	285	3,584
American	106,418	112,297	99,098	52,685	63,216	244,203	170,410	64,403	27,298	135,352
Arrow	762	2,038	1,445	902	4,230	8,990	1,083	2,223	408	6,980
Braniff	11,056	10,853	10,841	5,427	8,481	39,800	31,311	4,439	1,110	17,567
Continental	24,284	23,624	20,721	14,163	16,044	72,718	46,286	24,201	5,043	39,847
Delta	113,909	113,633	106,008	51,289	53,164	236,627	146,507	58,914	19,703	95,614
Eastern	98,051	95,837	96,533	47,651	66,854	225,567	122,658	57,693	23,368	103,643
Frontier	18,437	13,329	13,044	5,653	11,076	32,009	14,171	13,356	3,153	19,388
Hawaiian	3,144	3,019	2,419	1,235	2,540	6,074	2,511	2,838	505	4,550
Midway	1,800	5,212	2,098	1,415	3,285	9,326	2,889	3,352	1,424	8,614
New York Air	1,188	3,170	1,152	902	2,218	7,855	4,026	2,757	882	9,280
Northwest	27,844	32,879	30,887	15,744	15,232	96,182	53,146	28,384	11,564	30,370
Ozark	13,544	9,949	10,526	4,706	8,356	25,175	11,473	5,642	2,074	15,214
Pacific Southwest	11,123	9,696	11,307	5,752	7,643	25,346	16,237	12,038	2,190	18,627
Pan American	39,079	21,946	20,587	16,816	9,109	59,324	106,996	27,088	3,923	30,792
People Express	4,772	4,377	3,727	3,341	7,252	26,631	11,285	9,773	2,569	23,224
Piedmont	21,390	19,941	22,507	9,564	14,463	51,585	27,812	23,297	2,841	29,232
Republic	43,589	37,234	39,593	17,296	20,946	96,734	36,081	25,093	12,074	51,526
Southwest	7,699	8,001	7,831	5,185	6,320	29,426	11,676	12,407	1,273	13,152
Texas Intl	10,002	7,086	7,829	3,023	5,205	24,655	10,285	5,496	1,783	15,931
Trans World	47,410	56,045	57,357	40,457	44,232	133,974	106,868	42,212	17,590	59,359
Transamerica	1,671	1,935	8,385	2,956	8,592	16,575	2,645	4,513	1,052	8,502
United	104,150	122,467	141,899	80,965	93,096	307,844	225,501	78,663	28,368	143,744
USAir	36,973	31,499	40,210	16,567	23,889	73,870	39,449	20,740	9,517	44,294
Western	27,892	27,176	26,813	15,265	14,054	71,550	35,380	22,722	7,067	34,331
World	1,401	5,455	2,753	2,928	4,497	18,222	2,889	2,675	908	11,197
Industry Total	$ 30,311	$ 30,285	$ 30,532	$16,401	$ 19,872	$ 75,064	$ 48,007	$21,269	$ 7,276	$ 37,498

References

Baily, E. E., D. R. Graham, and D. P. Kaplan. 1985. *Deregulating the Airlines: An Economic Analysis*. Cambridge, MA: MIT Press.

Banker, R. D., G. Potter, and R. G. Schroeder. 1992. An empirical study of manufacturing overhead cost drivers. Working paper, University of Minnesota, Minneapolis, MN.

Belsley, D. A., E. Kuh, and R. S. Welsch. 1980. *Regression Diagnostics*. New York: John Wiley and Sons.

Borenstein, S. 1992. The evolution of U.S. airline competition. *Journal of Economic Perspectives* 6 (Spring): 45–73.

Breusch, T. S., and A. R. Pagan. 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica* 47 (September): 1287–94.

Caves, D. W., L. R. Christensen, and M. W. Tretheway. 1984. Economies of density versus economies of scale: Why trunk and local service airline costs differ. *Rand Journal of Economics* 15 (Winter): 471–89.

Christensen, L. R., and D. W. Jorgenson. 1969. The measurement of U.S. real capital input, 1929–1967. *The Review of Income and Wealth*, Series 15, 14 (December): 293–320.

Cooper, R., and R. S. Kaplan. 1987. How cost accounting systematically distorts product costs. In *Accounting and Management: Field Study Perspectives*, edited by Bruns and Kaplan. Boston: Harvard Business School Press.

Department of Transportation. 1990. *Secretary's Task Force on Competition in the U.S. Domestic Airline Industry*. Washington, D.C.: U.S. Government Printing Office.

Doran, H. E., and W. E. Griffiths. 1983. On the relative efficiency of estimators which include the initial observations in the estimation of seemingly unrelated regressions with first-order autoregressive disturbances. *Journal of Econometrics* 23 (October): 165–91.

Foster, G., and M. Gupta. 1990. Manufacturing overhead cost driver analysis. *Journal of Accounting & Economics* 12 (January): 309–37.

General Accounting Office. 1985. *Deregulation: Increased Competition is Making Airlines More Efficient and Responsive to Consumers*. Washington, D.C.: U.S. Government Printing Office.

———. 1990. *Airline Competition: Industry Operating and Marketing Practices Limit Market Entry*. Washington, D.C.: U.S. Government Printing Office.

Goldfeld, S. M., and R. E. Quandt. 1972. *Nonlinear Methods in Econometrics*. Amsterdam, The Netherlands: North-Holland.

Graham, D. R., and D. P. Kaplan. 1982. *Competition and the Airlines: An Evaluation of Deregulation*. Staff report, Office of Economic Analysis, U.S. Civil Aeronautics Board, Washington, D.C.

Hayes, R. H., and K. M. Clark. 1985. Explaining observed productivity differentials between plants: Implications for operations research. *Interfaces* 15 (November–December): 3–14.

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lutkepohl, and T. Lee. 1985. *The Theory and Practice of Econometrics*, 2d ed. New York: John Wiley and Sons.

Kirby, M. G. 1986. Airline economies of "scale" and Australian domestic air transport policy. *Journal of Transport Economics and Policy* 20 (September): 339–52.

Kmenta, J. 1971. *Elements of Econometrics*. New York: MacMillan.

———, and R. F. Gilbert. 1970. Estimation of seemingly unrelated regressions with autoregressive disturbances. *Journal of the American Statistical Association* 65 (March): 186–97.

Miller, J. G., and T. E. Vollmann. 1985. The hidden factory. *Harvard Business Review* 63 (September–October): 142–50.

Park, R. E., and B. M. Mitchell. 1980. Estimating the autocorrelated error model with trended data. *Journal of Econometrics* 13 (June): 185–201.

Parks, R. W. 1967. Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. *Journal of the American Statistical Association* 63 (June): 500–09.

Schiff, J. B., ed. 1991. *Cost Management Update*. Montvale, NJ: National Association of Accountants (now the Institute of Management Accounts) 1–2.

Sickles, R. C. 1985. A nonlinear multivariate error components analysis of technology and specific factor productivity growth with an application to the U.S. airlines. *Journal of Econometrics* 27 (January): 61–78.

———, and D. Good. 1986. Allocative distortions and the regulatory transition of the U.S. airline industry. *Journal of Econometrics* 33 (October–November): 143–63.

Strazheim, M. R. 1969. *The International Airline Industry*. Washington, D.C.: The Brookings Institute.

Theil, H. 1957. Specification errors and the estimation of economic relationships. *Review of the International Statistical Institute* 25: 41–51.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48 (May): 817–38.

White, L. J. 1979. Economies of scale and the question of "natural monopoly" in the airline industry. *Journal of Air Law and Commerce* 44: 545–73.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* 58 (December): 977–92.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 602–614

Simultaneous Estimation of Cost Drivers

Srikant M. Datar
Stanford University
Sunder Kekre
Tridas Mukhopadhyay
Kannan Srinivasan
Carnegie Mellon University

SYNOPSIS AND INTRODUCTION: Managers frequently choose the amounts to expend in various activities simultaneously rather than sequentially. Quality costs provide a common example. When managing quality, decisions to invest in different types of prevention activities are made jointly. For example, spending more on maintenance simultaneously reduced the spending necessary on supervision. Similarly, scrap costs are often traded off against the costs of prevention and appraisal activities.

Our article is motivated by field observations at an automobile lamp manufacturing plant. Specifically, we estimate two observed effects: (1) the influence of lamp design on the consumption of overhead resources during manufacturing (e.g., the effect of multicolor designs on supervision costs) and (2) the interdependence among supervision, maintenance, and scrap costs.

One way to understand cost drivers and manage costs is to employ an activity-based costing approach (Cooper and Kaplan 1991, chap. 5; Young and Selto 1991). With this approach, prevention costs of supervision and maintenance are allocated to products on the basis of hours of supervision and maintenance, and scrap costs are apportioned on the basis of physical scrap levels. Quality-related costs are reevaluated after products are redesigned and processes reconfigured to determine if quality related costs have indeed decreased. With such an approach, simultaneous effects of costs are not estimated.

We thank three anonymous reviewers for their comments and valuable suggestions. We are grateful for helpful discussions with Robert Kaplan, Allan Kleidon, Peter Reiss, and Seenu Srinivasan. The paper benefited from comments from seminar participants at the Harvard Business School, University of Colorado at Boulder, University of Michigan, and Vanderbilt University. The usual disclaimer applies.

Submitted August 1991.
Accepted February 1993.

In our approach, we simultaneously estimate interdependencies among activities. Instead of supervision hours, maintenance hours, and physical scrap levels, we use product and process design variables as cost drivers of supervision, maintenance, and scrap costs. Selecting product and process variables as cost drivers allows us to estimate the effect of alternative lamp designs on quality costs incurred during manufacturing. We also explicitly consider simultaneity. For example, our estimation procedure recognizes that maintenance costs affect supervision costs and vice versa and that both costs are affected by product and process design choices.

Our analysis provides valuable information to managers. At our site, designers use quality costs associated with different design features to guide future product designs and modifications. Similarly, as operations managers experiment with different methods to manage complexity, the simultaneous cost estimation enables them to evaluate which prevention activities are successful in reducing scrap.

Key Words: *Product costing, Cost drivers, Simultaneous cost estimation, Costs of quality, New products.*

Data Availability: *The data upon which this article is based may be obtained from the authors on request, subject to release by the company.*

THE organization of the paper is as follows. Section I describes the research site and provides a description of the cost categories and product and process characteristics that drive cost. Section II discusses the simultaneous equations estimation of our study. Key results are presented in section III. Concluding comments are presented in section IV.

I. Research Site

Our research site is a manufacturing facility for lamp assemblies for automobiles and trucks. A lamp has two components, a lens and a housing. But the products are diverse, ranging from simple lamps, such as park and signal lamps, to complex rear lamps equipped with multicolor lenses with special geometry and intricate optics. An injection-molding operation is the primary manufacturing process at the plant. The process consists of shooting molten plastic into a mold (also called a die or tool), which is designed to give the lens its desired shape and form when the plastic cools and solidifies.

Overhead costs at the plant are a substantial proportion of total costs, and are the focus of attention for cost management. We examine four major categories of overhead—supervision, tool maintenance, quality control and inspection, and scrap cost. The existing cost-accounting system assigns these overhead costs to products on the basis of traditional drivers, labor and machine hours. To verify whether the overhead costs allocated are representative of the demands placed on resources by individual products, we directly identify the overhead resources of supervision, inspection, tool maintenance, and scrap associated with individual products. At our research site, direct identification is possible due to the nature of production flows and the level of specialization at the plant. Each part is typically produced in a single molding department. Supervisors assigned to a specific department oversee the production of only 10–15 components.

Description of Costs

Supervision costs account for careful monitoring by supervisors to reduce tool breakages, tool maintenance, and the need for preventive maintenance. Part defects are also reduced as more care is taken by supervisors to prevent such defects from occurring. The time spent by supervisors on individual components was determined using structured interviews. The data are validated, triangulated, and cross-checked by using informal logs, sampling, and observation by the researchers.

Tool maintenance costs consist of the direct labor and material costs incurred in maintaining molds and the indirect costs of tool room supervision, stocking and tracking of tools, and maintenance of tool room equipment. Tool maintenance activities include cleaning the mold after use and stripping it after a certain number of cycles. Maintenance requires replacing moving parts (e.g., rollers, core lifts, and sliding cores), resurfacing the mold for surface finish, ensuring the proper and smooth functioning of moving parts, and making necessary adjustments whenever lens tolerances are violated. The tool maintenance resources consumed by individual products were determined from detailed toolroom records of the time spent on maintaining and repairing individual tools. Each tool is uniquely identified with a particular product so that tool maintenance can be directly traced to individual products.

Quality control and inspection costs are the indirect costs of quality. They include the efforts of technicians and inspectors responsible for ensuring that only defect-free products are sent out to assembly centers. As the probability of a product being out of tolerance increases, greater quality control costs are incurred to screen out defective units. An approach similar to supervision costs was adopted to collect data on quality control and inspection costs.

Scrap costs are computed on the basis of detailed records of rejected lenses and revised identification of costs to products. The quantity of rejects recorded by supervisors was validated with records maintained by the quality assurance department.

The correlations among the per unit costs of supervision, tool maintenance, quality control and inspection, and scrap across various products (computed on the basis of the time spent and resources consumed) are positive (see table 1). However, as discussed in section III, after controlling for complexity, the correlations between prevention activities (supervision and tool maintenance) and scrap costs become negative.

Factors of Product and Process Complexity

Our choices of complexity factors that drive demands for overhead resources were based on detailed studies of the plant, and knowledge of the physics and engineering aspects of the plastic injection-molding process. The critical activity in the molding process is the functioning of the mold or tool itself. The complexity of the mold increases with the cross sectional area and the number of moving parts required to give the lens its shape, contours, and curves. The area over which the molten plastic flows

Table 1
Correlations Among Cost Categories

Variables	*Supervision*	*Tool Maintenance*	*Quality Control*	*Scrap*
Supervision	1.000	0.280	0.697	0.245
Tool Maintenance	0.280	1.000	0.124	0.186
Quality Control	0.697	0.124	1.000	0.232
Scrap	0.245	0.186	0.232	1.000

and cools is a measure of complexities resulting from the size of the mold as well as manufacturing complications associated with heating and cooling cycles. A larger area increases the problems of spots, cracks, and "short shots" (which arise when the entire mold is not filled with molten plastic). Hence, we define moving-part complexity as the product of the number of moving parts and the cross-sectional area.[1] Data on moving parts and area were obtained from the tool specification report and were cross-checked with the tool audit report.

Multicolor molding is a complex operation requiring molten plastic of two or more colors to be injected into the mold at different points in the manufacturing cycle. The complexity arises because, to prevent two colors from mixing, the first color must completely dry out and solidify before the next color is injected. The process requires careful and rigorous monitoring of temperatures and pressures, and heating and cooling cycles. The advantage of multicolor molding is the excellent finish imparted to the lens. The complexity of multicolor molding increases with the surface area over which the plastic must flow and form. We therefore define multicolor molding complexity as the product of a dummy variable (1 if the product is multicolor molded, 0 otherwise) and area. Data on multicolor molded parts and area were obtained from the product design records and were cross-checked with the tool specification report.

Supervisors have to ensure that the temperature profile during start-up follows specifications, and molding of production parts begins only after the required thermal stability is achieved. The duration of the warm-up period varies by products and machine instrumentation. The longer the period, the greater the care and attention required from supervisors during production runs. The thermal stability factor, defined as the ratio of the run length to warm-up period, captures the relative complexity of the molding process from a supervisory standpoint. This factor was obtained from the process specification drawn up by the quality assurance department.

Depth captures the complexity of plastic flow to mold a part. Also, deep tools are difficult to maintain and modify since certain parts of the mold are harder to reach and clean. Furthermore, the quantity of plastic required across the length of the lens is not uniform. This causes complications in the heating and cooling cycles. The depth of each component was measured in inches and obtained from product design and tool specification reports.

---

[1] This measure is closely correlated with the number of moving parts in the mold. The interaction term is a better representation of engineering complexity.

Table 2
Summary Statistics
($N=121$)

Cost Variable	Mean	Variance
Endogenous (cost per unit)		
Supervision	0.056	0.036
Tool maintenance	0.171	0.049
Quality control	0.007	0.000
Scrap	0.510	0.864
Exogenous		
Direct labor hours (per unit)	0.826	1.622
Moving-parts complexity (per unit)	0.006	0.000
Multicolor molding	4.300	5.782
Thermal factor	8.245	3.178
Number of functions	1.917	1.360
Machine complexity	0.157	0.262
Depth	1.092	1.198
Direct labor inspection (hours per unit)	0.050	0.002

Lenses frequently perform multiple functions. For example, a multifunction (or combination) lamp may have a rear lamp, license plate lamp, and back-up lamp all combined into a single product. These lamps must be built to meet the functional specification of each component product. Combination lamps require simultaneous satisfaction of tolerances and complex interactions among tolerances, and hence place increased demand on manufacturing overhead resources. The variable for number of functions measures the number of different lamps that are combined into a single lamp. Data on number of functions were obtained from detailed part drawings.

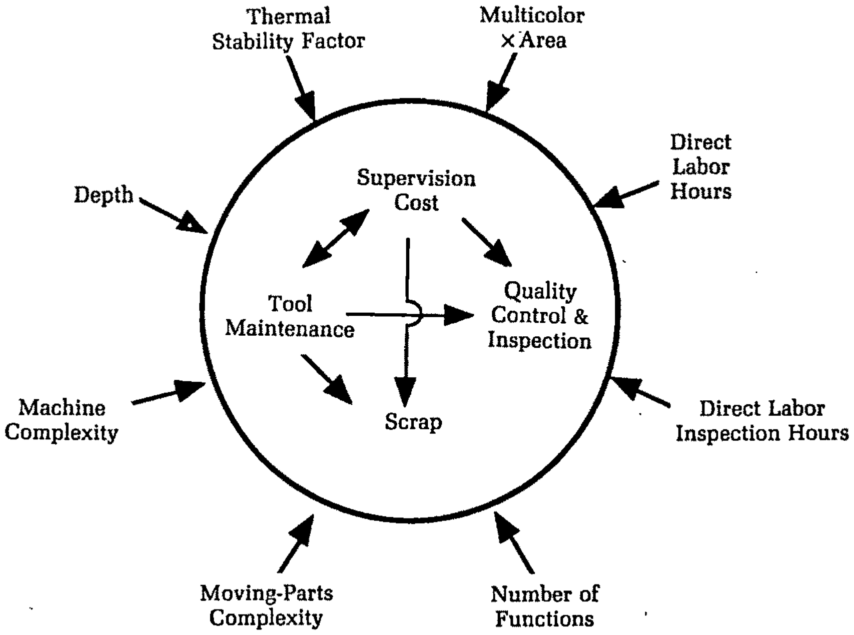
Complex machines, by definition, are difficult to operate. These machines typically have multiple controls that need to be properly synchronized for the production of good components. High-tonnage machines are also difficult to operate. The large force applied by large machines demands that the mold be carefully set to prevent damage to it and to ensure the production of defect-free units. We use the replacement cost of the machines as a surrogate for complexity and size of the machines. This information was obtained from the accounting and finance departments.

Summary statistics for all the key variables are presented in table 2. The variables are categorized as either endogenous or exogenous, which will be explained in the next section.

II. Simultaneous Estimation

As discussed earlier, product and process features committed at the design stage (e.g., moving parts in the mold and multicolor molding), affect the consumption of overhead resources during manufacturing. Further, interdependencies exist among cost categories. The thrust of our empirical estimation is to recognize that various activities are simultaneously related and do not arise sequentially or independently. For example, both product and process complexity factors and resources expended on prevention and appraisal drive scrap costs.

Figure 1
Proposed Simultaneous Cost-Engineering System



Note: The exogenous variables are represented outside the enclosed endogenous variables.

We attempt to capture the effects of complexity factors as well as other managerially controllable activities on the consumption of various overhead resources. We refer to the cost drivers of product design and process engineering (such as characteristics of the mold and the type of equipment) as exogenous variables (see also Banker et al. 1990) because these variables are determined outside the plant domain. Variables (such as the shape, size, and contours of individual lamps) are chosen by assembly divisions on the basis of consumer preferences, aesthetics, and functionality. Exogenous variables affect variations in resources of supervision, tool maintenance, quality control, and scrap according to variations in the manufacture of individual lamps within the plant.

We classify supervision, tool maintenance, quality control, and scrap resources as *endogenous* because their values are determined simultaneously within the plant in response to exogenous factors. The interactions among endogenous variables determine the econometric methodology to estimate the drivers of supervision, tool maintenance and quality control activities, and resources lost by way of scrap. We estimate the resources consumed within the plant as a system of simultaneous equations. Failure to recognize interdependencies among resources leads to erroneous estimates and incorrect conclusions about the effects of exogenous and endogenous variables on activities.

A schematic representation of the interdependencies among variables is presented in figure 1. All exogenous variables are outside the encircled endogenous variables.

Each endogenous variable is determined by the effect of a set of influencing exogenous variables and simultaneity effects from other endogenous variables.

We estimate the following system of equations where the endogenous variables are: *SUPCOST/UNIT*, the supervision cost of each lens; *TMCOST/UNIT*, the cost in tool maintenance incurred per unit; *QCCOST/UNIT* the cost per unit of quality control and inspection; and *SCRAPCOST/UNIT*, the cost of scrap generated per unit.

$$\begin{aligned} SUPCOST/UNIT = {} & \alpha_{11} + \beta_{11}(MVGPTS\ COMPL)/UNIT + \beta_{12}(MULTCLR*AREA) \\ & + \beta_{13}THERMFAC + \beta_{14}NUMFUNC + \beta_{15}MACHCOMPL \\ & + \beta_{16}DIRLABHRS/UNIT + \gamma_{11}TMCOST/UNIT + \epsilon_1. \end{aligned} \tag{1}$$

$$\begin{aligned} TMCOST/UNIT = {} & \alpha_{21} + \beta_{21}(MVGPTS\ COMPL)/UNIT + \beta_{22}(MULTCLR*AREA) \\ & + \beta_{23}NUMFUNC + \beta_{24}DEPTH + \beta_{25}MACHCOMPL \\ & + \gamma_{21}SUPCOST/UNIT + \epsilon_2. \end{aligned} \tag{2}$$

$$\begin{aligned} QCCOST/UNIT = {} & \alpha_{31} + \beta_{31}(MVGPTS\ COMPL)/UNIT + \beta_{32}(MULTCLR*AREA) \\ & + \beta_{33}DEPTH + \beta_{34}MACHCOMPL + \beta_{35}DINSPHRS/UNIT \\ & + \gamma_{31}SUPCOST/UNIT + \gamma_{32}TMCOST/UNIT + \epsilon_3. \end{aligned} \tag{3}$$

$$\begin{aligned} SCRAPCOST/UNIT = {} & \alpha_{41} + \beta_{41}(MVGPTS\ COMPL)/UNIT + \beta_{42}(MULTCLR*AREA) \\ & + \beta_{43}NUMFUNC + \beta_{44}DEPTH + \beta_{45}MACHCOMPL \\ & + \gamma_{41}SUPCOST/UNIT + \gamma_{42}TMCOST/UNITF + \epsilon_4. \end{aligned} \tag{4}$$

The exogenous variables are described as follows:

DEPTH = depth of a lens in inches,
DINSHRS/UNIT = measure of the direct labor inspection hours per unit, as determined by the standard operating procedures for the product,
DIRLABHRS/UNIT = measure of the direct labor hours per unit based on process-routing sheets,
MACHCOMPL = machine complexity measured by the replacement cost of each machine,
MULTCLR×AREA = the area of a multicolor mold,
MVGPTS COMPL/UNIT = the complexity of the sweep path of the moving parts in the mold. We divide by production volume since overhead costs are incurred setting up the mold, independently of the volume of production,
NUMFUNC = the number of functions in the lens, and
THERMFAC = the thermal stability factor associated with each lens and is measured as the ratio of the run period to the warm-up period.

The system of equations describes interactions among resources consumed. We expect the sign of the coefficient (β_{ij}) of each exogenous variable (except the thermal stablity factor) in the system of equations to be positive. An increase in complexity factors increases the demand on overhead resources. We elaborate next on the endogenous variables in the system of equations.

The first equation describes the drivers of supervisory resources. Tool maintenance cost per unit is an explanatory variable in equation (1). Fault-free functioning of the tool is a critical element in the smooth and continuous production of lens components. The effective maintenance of tools reduces the need for adjustments and reruns, and thereby reduces supervisory costs. We hypothesize that the coefficient of tool maintenance cost per unit (γ_{11}) will be negative in equation (1).

The second structural equation reflects the expectation that additional resources devoted to supervision have a beneficial effect on tool maintenance costs. Careful monitoring and supervision and better handling of the tool ensure proper alignment and functioning of moving parts. Thus tools remain within tolerance limits for longer periods, reducing the demand for tool maintenance. We hypothesize that the coefficient of supervisory cost per unit (γ_{21}) will be negative in the second structural equation (2), which suggests a substitution effect.

We specify quality control and inspection costs per unit as a derived endogenous variable. That is, the amount of quality control and inspection expenditure on a product depends on the firm's experience with defect rates for that product. If defect rates for a product are high, greater overhead resources of quality control and inspection are expended on that product. Direct inspection hours reflect the intrinsic complexity of line inspection demanded by a product from line operators.

Equation (3) specifies that resources devoted to prevention activities of tool maintenance and supervision decrease quality control costs per unit. We therefore hypothesize that the sign of the coefficient for tool maintenance cost per unit (γ_{32}) and that for supervisory cost per unit (γ_{31}) are negative. Quality control and inspection costs have no influence, however, on supervision and tool maintenance costs. Quality assurance activities at this site are reactive in nature. The major objective of inspection is to identify defective products before faulty components are shipped to assembly plants.

Similarly, the structural equation for scrap specifies the effect of supervision and tool maintenance on scrap costs. The signs of the coefficients of supervisory cost per unit (γ_{41}) and tool maintenance cost per unit (γ_{42}) are anticipated to be negative in equation (4). The equation reflects the widely held perception that increased prevention activities improve quality and mitigate failure. The implication for cost driver analysis is to recognize that costs incurred in prevention (tool maintenance and supervision) are drivers of scrap costs.

Ordinary least squares (OLS) cannot be applied on an equation-by-equation basis in this case because the disturbance term and endogenous explanatory variables are correlated. The direct application of least squares to each equation yields biased and inconsistent estimates of the coefficients. The system of equations are estimated by using the two-stage least squares procedure. The two-stage technique yields unbiased and consistent parameters. The structural form of the system of equations reveals direct effects of complexity and the endogenous relationships among overhead resource categories. The reduced form is easily obtained by solving the system of simultaneous equations for the endogenous variables.

III. Results and Discussion

The results[2] of the two-stage least squares procedure for the structural equations are presented in table 3 and OLS results are reported in table 4. As noted earlier, OLS estimates are biased and inconsistent. For example, OLS estimates lead to the erroneous conclusion that moving-parts and machine complexity have a significantly lower effect on unit supervision cost.[3] The results with two-stage least squares show the effect of moving-parts and machine complexity to be significantly greater.[4]

---

[2] The parameter estimates are very robust. We reestimated the model after dropping 5 percent of the observations at random. All the coefficients and their signficance levels were found to be stable. We also checked for multicollinearity using the Belsey et al. test (1980).

The structural equations characterize two effects. The first is the direct effect of exogenous variables such as moving parts, multicolor molding, thermal stability, number of functions, depth, and machine complexity on supervision, tool maintenance, quality control and inspection, and scrap resources. For example, a unit increase in moving-parts complexity results in expected direct effects of $16.26 on unit supervision cost (see table 3, panel A), $20.33 on unit tool maintenance cost (panel B), $7.97 on unit quality control cost (panel C) and $24.27 on unit scrap cost (panel D). Our results suggest that products and process design features, committed at the design stage, significantly affect manufacturing overhead costs at our site. The cost effects of specific features on subsequent quality-related manufacturing overheads constitute useful information for designing new products and modifying existing ones. For example, managers and designers at our site have begun focusing on designing fewer moving parts into lamps as a means of managing costs.

The second effect is the indirect influence of endogenous variables. The coefficient estimates for the endogenous variables in the structural equations in the second stage of the two-stage procedure (reported in table 3) are all negative and significant.[5] Prevention activities are chosen simultaneously and reduce scrap costs at our site. In particular, increasing supervisory resources, in response to exogenous increases in complexity, by $1 per unit has the effect of decreasing tool maintenance cost by $0.20 per unit (see panel B of table 3), quality control cost by $0.07 per unit (panel C), and scrap cost by $0.30 per unit (panel D).

The net effect can be derived by simultaneously solving the system of equations and describing each overhead cost (endogenous variable) in terms of exogenous factors alone. The derived system of equations are the structural equation in the reduced form. For example, the reduced form coefficients of moving-parts complexity (the effect of a unit increase in moving-parts complexity on each overhead cost category) is as follows:

Supervisory cost per unit	$12.18
Tool maintenance cost per unit	$17.93
Quality control inspection cost per unit	$ 3.24
Scrap cost per unit	$16.91

Note that the corresponding OLS coefficients in table 4 differ substantially (−57 percent to +12 percent) in magnitude as compared to the reduced form results. Also, the net effect of an increase in moving-parts complexity on each category of cost is less than the direct effect of complexity on each cost individually from the results in table 3. The structural equations indicate that this lower net effect is attributable to favorable interaction among cost categories. For example, the direct effects of moving-part com-

---

[3] In fact, the OLS estimates suggest that higher machine complexity (−0.025) lowers supervision cost.

[4] We also estimate the system of equations using three-stage least squares. Our results are robust to these estimation procedures.

[5] We confirmed the endogenous nature of *SUPCOST* and *TMCOST* by conducting the test proposed by Spencer and Berk (1981).

Table 3

Two-Stage Least Squares Estimates

Panel A. Dependent Variable, Supervision (Cost / Unit):

Independent Variable	*Coefficient*	*Standard Error*	T-*Ratio*
Intercept	0.0016	0.0155	0.109
Moving Parts Complexity / Unit	16.2590	2.1791	7.461
Multicolor × Area	0.0034	0.0009	3.756
Thermal Stability Factor	−0.0011	0.0001	−8.179
Number of Functions	0.0004	0.0074	0.064
Machine Complexity	0.3167	0.0584	5.424
Direct Labor Hours / Unit	0.0314	0.0123	2.548
Tool Maintenance Cost / Unit	−0.2274	0.0845	−2.691

Panel B. Dependent Variable, Tool Maintenance (Cost / Unit):

Independent Variable	*Coefficient*	*Standard Error*	T-*Ratio*
Intercept	0.1190	0.0323	3.673
Moving Parts Complexity / Unit	20.3288	5.0001	4.066
Multicolor × Area	0.0041	0.0025	1.619
Number of Functions	−0.0105	0.0178	−0.591
Depth	0.0796	0.0240	3.309
Machine Complexity	0.1961	0.0970	2.021
Supervision Cost / Unit	−0.1967	0.0540	−3.637

Panel C. Dependent Variable, Quality Control (Cost / Unit):

Independent Variable	*Coefficient*	*Standard Error*	T-*Ratio*
Intercept	0.1518	0.0034	44.226
Moving Parts Complexity / Unit	7.9734	0.7666	10.400
Multicolor × Area	0.0009	0.0003	3.114
Depth	0.0214	0.0029	7.360
Machine Complexity	0.0373	0.0117	3.174
Direct Inspection Hours / Unit	0.2552	0.1040	2.453
Supervision Cost / Unit	−0.0761	0.0239	−3.182
Tool Maintenance Cost / Unit	−0.2120	0.0294	−7.191

Panel D. Dependent Variable, Scrap (Cost / Unit):

Independent Variable	*Coefficient*	*Standard Error*	T-*Ratio*
Intercept	0.4780	0.0333	14.323
Moving Parts Complexity / Unit	24.2689	4.6384	5.232
Multicolor × Area	0.0059	0.0022	2.636
Number of Functions	0.0892	0.0147	6.319
Depth	0.0736	0.0195	3.774
Machine Complexity	0.4497	0.0807	5.571
Supervision Cost / Unit	−0.2994	0.1849	−1.619
Tool Maintenance Cost / Unit	−0.2070	0.1178	−1.758

Table 4
OLS Estimates Using Only Exogenous Variables

Panel A. Dependent Variable, Supervision (Cost / Unit; $R^2=0.84$*):*

Independent Variable	*Coefficient*	*Standard Error*	*T-Ratio*
Intercept	−0.0097	0.0141	−0.690
Moving Parts Complexity / Unit	11.7702	2.0321	5.811
Multicolor × Area	0.0071	0.0005	16.751
Thermal Stability Factor	−0.0008	0.0001	−10.933
Number of Functions	0.0403	0.0534	0.754
Machine Complexity	−0.0251	0.0089	−2.823
Direct Labor Hours / Unit	1.6065	0.8138	1.973

Panel B. Dependent Variable, Tool Maintenance (Cost / Unit; $R^2=0.39$*):*

Independent Variable	*Coefficient*	*Standard Error*	*T-Ratio*
Intercept	0.1141	0.0323	3.523
Moving Parts Complexity / Unit	13.9803	4.6694	2.994
Multicolor × Area	0.0021	0.0019	1.073
Number of Functions	0.0174	0.0119	1.457
Depth	−0.0808	0.0242	−3.341
Machine Complexity	0.3035	0.0922	3.291

Panel C. Dependent Variable, Quality Control (Cost / Unit; $R^2=0.62$*):*

Independent Variable	*Coefficient*	*Standard Error*	*T-Ratio*
Intercept	0.0292	0.0045	6.521
Moving Parts Complexity / Unit	1.4002	0.6440	2.174
Multicolor × Area	−0.0001	0.0002	−0.460
Number of Functions	0.0041	0.0025	1.602
Machine Complexity	0.0114	0.0137	0.832
Direct Inspection Hours / Unit	0.4626	0.2637	1.754

Panel D. Dependent Variable, Scrap (Cost / Unit; $R^2=0.76$*):*

Independent Variable	*Coefficient*	*Standard Error*	*T-Ratio*
Intercept	0.4529	0.0285	15.891
Moving Parts Complexity / Unit	18.8597	3.8004	4.963
Multicolor × Area	0.0037	0.0017	2.189
Number of Functions	0.0917	0.0142	6.465
Depth	0.0587	0.0182	3.224
Machine Complexity	0.4365	0.0770	5.669

plexity on quality control and scrap costs are $7.97 (panel C of table 3) and $24.27 (panel D), respectively. The increased prevention expenses on supervision and tool maintenance, however, restrict the net effect of moving part complexity on quality control to $3.24 and to $16.91 on scrap.

Quality improvement programs focus on understanding the causes of scrap and reducing defects generated in the plants. Our two SLS analysis indicates that the scrap costs are significantly influenced by underlying complexity drivers. But the adverse effect of increased complexity on scrap costs is mitigated by prevention activities and better management of complexity through tool maintenance and supervision activities. The two-stage approach offers a better understanding of the underlying structure with respect to the effects of complexity and the payoffs from prevention activities. These insights are not available from OLS or reduced form analysis.

Operations managers at our site handle various types of complexity differently. The reduced form suggests that managers respond to moving-part complexity by investing more in tool maintenance than in supervision.[6] Conversely, they respond to machine complexity by investing in greater supervision rather than tool maintenance.[7]

Note that the above discussion is based on the current operating practices and may not reflect optimal resource allocations. The analysis, however, provides guidance for managing product features to lower costs, and improve profitability. Further insights on managing complexity can be gained by replicating our analysis over time across plants or examining sites that choose different responses to complexity. The effects on failure and scrap can then be compared by using the two-stage method that controls for differences in complexity across the plants.

Our results also provide empirical support for the ABC hypothesis that non-volume drivers (i.e., drivers other than direct labor and machine hours) affect the demand for overhead resources. At our site, for example, direct labor explains only a fraction of supervisory overhead. We find that cost drivers such as moving parts, multicolor molding, and machine complexity significantly affect overhead resources.

IV. Conclusion

The contribution of our analysis in understanding cost structures is two-fold. First, we recognize and incorporate simultaneity among overhead costs in estimating drivers of these costs. Second, we measure the effect of product and process features on overhead costs in the presence of simultaneity. As demonstrated earlier, failure to recognize simultaneity results in inaccurate estimates of the effect of the cost drivers.

The analysis yields information for making design modifications by quantifying the manufacturing cost impact of product features. For example, any design changes that reduce the number of moving parts in the mold substantially lowers the expenditures in supervision, tool maintenance, and quality control. The estimates from the structural equations provides insights into the direct effect of complexity and reveal the interdependencies between cost categories. The reduced form estimates yield the net effect. These estimates may mask the adverse direct effect of complexity due to substitution between cost categories. Operations managers can utilize the information from the estimates of the structural equations to make appropriate resource allocations to manage complexity.

[6] The net effect, as obtained from the reduced-form equation, of a unit increase in moving-part complexity is $12.18 on supervision cost and $17.93 on tool maintenance cost.

[7] In this case, the reduced-form results indicate that a unit increase in machine complexity affects supervision cost by $0.28 and tool maintenance by $0.14.

We believe that the investigation of simultaneity among cost categories offers scope for future work. Our effort is a first attempt to understand trade-offs among activities. Additional analyses and replication at different sites are required to gain a comprehensive understanding of the choice of cost drivers and the effect on scrap costs of different prevention strategies. Further research is necessary to establish better ways to estimate production functions and manage cost drivers and simultaneous interactions among costs.

References

Banker, R., S. Datar, S. Kekre, and T. Mukhopadhyay. 1990. Costs of product and process complexity. In *Measures for Manufacturing Excellence*, edited by R. S. Kaplan. Boston: Harvard Business School Press, 269–90.

Belsey, D. A., E. Kuh, and R. S. Welsh. 1980. *Regression Diagnostics*. New York: John Wiley and Sons.

Cooper, R., and R. S. Kaplan. 1991. *The Design of Cost Management Systems*. Englewood Cliffs, NJ: Prentice-Hall.

Spencer, D. E., and K. N. Berk. 1981. A limited information specification test. *Econometrica* 49 (7): 1079–85.

Young, M. S., and F. H. Selto. 1991. New manufacturing practices and cost management: A review of the literature and directions for research. *Journal of Accounting Literature* 10: 265–98.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 615–620

A Perspective on Cost Drivers

Nicholas Dopuch
Washington University

SYNOPSIS AND INTRODUCTION: My task is to provide a perspective regarding the general topic of choosing optimal cost drivers, as that topic is reflected in the selected articles of this Forum. This perspective can be developed more easily if I first digress back to the 1960s when managerial accounting research was beginning to evolve to its present stages.

In the 1960s, teaching and research in managerial accounting were heavily influenced by the growing popularity of operations research and management science approaches to business problems. These approaches were largely normative in nature, consisting basically of optimization techniques that, if implemented, promised to improve managerial decision making. Their normative character was carried over into managerial accounting as well, with the basic theme of many managerial accounting articles in that era being, "Here is how optimization techniques may be used to guide the design of improved managerial accounting systems."[1] Rarely, if ever, did the researcher provide empirical evidence or other forms of proof that the new systems would in fact lead to improvements in managerial decisions. Rather, the main result provided was that the new systems could generate new accounting numbers under certain assumed conditions. If these new numbers appeared to be materially different from the old, implementation of the new systems could lead to changes in managers' decisions. The changes in managers' decisions were assumed to be desirable ones to induce. However, desirability was based primarily on subjective criteria adopted by the researcher.

Generally, a necessary condition for potential improvements in managerial decisions from new accounting systems is that the new system generate accounting numbers that are materially different from those obtained from the existing system. For this reason, I label such studies as materiality studies, since they basically follow the traditional materiality rule in accounting. But different accounting numbers might not lead to changes in decisions that result in payoffs sufficient to justify the costs of implementing the new systems. Indeed, recognition of this basic limitation of materiality studies was a strong motivation for managerial accounting researchers to investigate the advantages of using the information eco-

---

[1] A typical example is Dopuch et al. (1967); see also the cost allocation articles appearing about then.

Invited by the Editor.

nomics paradigm to estimate the net payoffs from implementing new accounting systems.

Recall that information economics is a conceptual framework that can be used to assess how new information will affect managerial decisions and, assuming decisions will change, the expected payoffs from implementing a new accounting information system. Included in the framework is an assumed decision-maker with a particular utility function who processes information through a specific decision model. The model, in turn, yields an *ex ante* optimal decision. The value of new information is then measured as the change in the net payoffs induced by moving from a previous optimal decision to a revised optimal decision under the assumption that the new information does in fact induce a revision in the optimal decision. Note that the original optimal decision might still remain optimal even though the new information received by the decision maker differs from the old.[2] Obviously, the conclusion in the latter case is that the new information system would have either a zero or negative expected net value.

I. Articles in This Forum

NOW back to the task at hand. In one of the articles in this Forum, Banker and Johnston estimate the relationships between various operations-based and volume-based cost drivers and the indirect costs incurred by U.S. airlines, using data that are in the public domain. Their purpose is to determine whether their selected cost drivers could be useful in analyses of overhead costs for making operating and investment decisions by the airlines. They find that their drivers are significantly associated with various categories of overhead costs. This leads them to conclude that their "empirical cost driver analysis is managerially significant for the industry and period that we have chosen to examine.... During the period of transition following deregulation [of the airline industry], carriers adopted a rich variety of strategies to improve productivity and increase market share. These strategies directly involved both volume- and operations-based cost drivers."

Banker and Johnston do not indicate which drivers the airlines actually used in developing their operating and investment strategies. Indeed, it is conceivable that the airlines are presently estimating costs using either the same drivers identified by Banker and Johnston or drivers highly correlated with these. In this case, the use of Banker and Johnston's drivers would not produce materially different accounting numbers from what the airlines are presently obtaining. Assuming that the operating and investment strategies of the airlines were based on cost estimates obtained from different sets of cost drivers, and that the Banker and Johnston drivers would produce materially different accounting numbers, one still would not know whether the airlines would have been better or worse off with the Banker and Johnston drivers. The recent failures and operating losses experienced by the major airlines do suggest, however, that whatever cost analyses were performed, the subsequent strategies adopted have not paid off that well.

---

[2] For example, new information may yield different contribution margins in a linear program, but the revised margins may not change the optimal decision.

In summary, Banker and Johnston show that potentially different cost estimation functions can be derived using a general activity-based costing analysis. I view their work as falling within the basic category of materiality studies as identified above. This means that it is not possible to ascertain whether their analyses provide more or less accurate cost estimates (the true costs faced by the airline industry are not known[3]) nor whether their resulting cost numbers are more or less useful than other alternatives, at least from the standpoint of an information economics perspective.

In the article by Datar et al. (1993), it is argued that, in the presence of simultaneous relations between activities and their costs, better estimates of the costs of the individual activities can be obtained by using two-stage regression estimation procedures rather than ordinary least squares. "Better" is defined relative to statistical criteria; that is the estimates of coefficients obtained from the two-stage regression are less biased and are consistent (converge to a mean value) in the presence of simultaneity of relations. Their results indicate that the two different estimation approaches generate different estimates of the regression coefficients.

If one accepts the statistical criteria of Datar et al. for assessing cost estimates, the two-stage estimates may be judged to be better. But again, in the absence of any explicit determination of how their estimates can actually affect operating personnel's decisions, the net economic benefits of any improvements in the cost estimates brought about by a two-stage estimation approach cannot be assessed. The conclusion is that, since the cost estimates under the two approaches are materially different, they could lead to different (and potentially improved) managerial decisions. More on this later.

The article by Babad and Balachandra attempts to go one step beyond a materiality approach by basically employing an information economics framework in their evaluation of different accounting (cost allocation) systems. The information economics framework is disguised in their presentation since they do not explicitly rely on a utility function, nor do they trace their different cost allocation numbers through a formal decision model to determine how optimal decisions change in response to different cost allocations. Nevertheless, their adoption of a squared error measure is consistent with a quadratic loss function for allocation errors, and their use of managements' weights to rank the severity of the magnitudes of squared errors is consistent with an information economics framework. More specifically, their use of different weights implies that squared errors of the same magnitude may have differential economic penalties.[4]

The materiality approach employed by Banker and Johnston and by Datar et al. avoid explicit specifications of how different cost allocations are used in particular decision contexts by individual decision-makers. Their implicit assumption is that users of the cost accounting numbers will prefer information systems which produce numbers which more closely adhere to some general criteria over systems which produce numbers which are further removed from the assumed criteria. But the relative net economic benefits which can be expected from implementing alternative accounting systems are ignored.

---

[3] Indeed, in a recent paper, Datar and Gupta (1992) argue that moving to more cost pools and selecting more cost drivers can result in greater total error in product costs.

[4] The loss equation used by Babad and Balachandran is similar to that found in Hwang et al. (1992) who base their loss function on the opportunity cost model found in Vives (1990). That model basically shows the opportunity cost incurred if a (single-product) firm deviates from its optimal quantity of output. For example, in the present context, this could happen if the costs allocated to individual products lead to incorrect quantity output decisions.

Of course, these authors are justified to some extent in following a materiality approach since it is commonly used by accountants when attempting to assess alternative accounting numbers that may be used by multiple parties for multiple purposes. Nevertheless, their analyses do not extend to the issue of how much error or inefficiency in cost estimated can and cannot be tolerated by decision-makers, given the net benefits from implementing different systems for different purposes. Moreover, no materiality study can demonstrate that one set of cost estimates is more accurate than another unless the researcher knows the true cost function (see Demski and Christensen 1992). This is rarely, if ever, known in practice. The information economics approach used by Babad and Balachandra has a stronger conceptual appeal, but it is susceptible to the fundamental criticism that it is difficult, perhaps impossible, to implement in practice. This becomes apparent if we consider several assumptions made by Babad and Balachandra to make their approach operational. First, they assume that the accountant knows the true cost drivers of activities and functional relationships tying the drivers to the total costs of activities. This permits them to measure the squared errors of cost estimates if other than these true drivers are used in the estimates. But if accountants possessed this knowledge, the issue of whether they should use all of the drivers or some optimal subset would seem to fade into triviality.

Second, Babad and Balachandra assume that operating managers can supply weights that allow the accountant to rank the severity of errors associated with alternative cost estimates which do not use the optimal drivers. Managers could provide these weights only if they were able to articulate exactly how costs are used in their decisions and, therefore, how different costing numbers will affect their *ex ante* optimal operating and investment decisions. The assumption that managers possess such knowledge seems totally unrealistic, to say the least.

It is not necessary to dwell further on the fundamental problems of trying to implement an information economics approach in the evaluation of alternative cost accounting systems. These problems were identified in the 1970s, and the inability to circumvent them eventually led researchers to conclude that informational economics provides a useful conceptual framework for structuring our thinking about alternative costing systems, but the paradigm can rarely be applied in actual evaluations of the relative desirability of different accounting systems.

Managerial accounting researchers' response to this practical dilemma in the 1970s was to embrace agency theory as an alternative research paradigm. In the process, they largely abandoned a normative approach to managerial accounting research. I believe one striking difference between activity-based costing, with its emphasis on identifying more appropriate cost drivers, and current agency research in managerial accounting is the avowed purpose of the former to provide more useful managerial accounting information. In contrast, the highly abstract formal analyses in agency research rarely lead to the derivation of empirical implications—either normative or positive. I do not know if analytical (agency) researchers in managerial accounting deliberately divorce their work from the real world, or if the divorce has evolved because of some underlying form of incompatibility.

II. Summary and Conclusions

Although managerial research based on agency theory may seem too far removed from practice, certain researchers who have been wooed over to activity-based costing may be moving too far in the direction of deriving normative implications from their managerial research. It seems that a number of activity-based costing studies are motivated simply to provide "better" accounting numbers to managers than what they have been receiving in the past. More important, "better" is often assumed to follow automatically from the premise that a selection of additional cost drivers and overhead cost pools will move researchers closer to the "true" cost numbers. This is not the kind of perspective that should guide serious research in the area; rather, it is a perspective one might expect to encounter in a consulting engagement.

One common deficiency I see in some of the activity-based costing studies of today is the failure to identify and then address a serious information issue and, once identified, to derive criteria which can be used to address it. With apologies to Banker and Johnston, I had difficulty identifying a particularly interesting issue pursued in their study. My best guess is that their basic intent was to identify the underlying activities of airlines' cost functions. In that sense, their study parallels the one by Datar et al., which basically represents an attempt to better understand the cost function(s) underlying the quality control activities of a specific firm. Under the assumption that certain subactivities within a quality control program are interrelated and that different trade-offs are possible within that program, their objective of trying to obtain better estimates of the costs associated with the different trade-offs would seem to be a useful one to pursue.

Banker and Johnston adopt a similar assumption regarding the kinds of trade-offs faced by the airlines. However, they do not provide any criteria for evaluating whether their resulting cost estimates dominate alternative estimates that could be derived, or that are actually being used by the airlines. In contrast, if one accepts the statistical criteria Datar and Gupta adopt to assess alternative estimates of quality control costs, their two-stage regression estimates of the marginal costs of the different subactivities can be regarded as more reliable than those obtained through ordinary least squares. Datar and Gupta state that the operating personnel in the quality control program of their sample firm were inclined to accept their estimates as more reliable than what they had been using. Nevertheless, they did not go the additional step and provide evidence that the operating personnel actually changed their quality control decisions and, if so, the extent to which the changes were beneficial.

Finally, I want to emphasize that I have no fundamental quarrel with Babad and Balachandran's decision to adopt an information economics approach to derive the optimal set of cost drivers in a particular costing situation. After all, that approach does incorporate criteria by which alternative cost estimates can be ranked. But the information economics approach was not found to be practical in the 1970s, and I am convinced similar kinds of impasses will be encountered in the 1990s if that paradigm is again adopted as a general approach to evaluations of substantive costing issues. For that reason, I would not encourage its reincarnation in managerial accounting research.

Surely there are some interesting cost estimation issues that can be addressed in field and analytical studies without requiring the researcher to either adopt a full-fledged information economics approach or identify the true cost function of an activity or product or service. One possibility is to use various statistical criteria that have general acceptance in other research areas to evaluate costing alternatives. Indeed, Babad and Balachandran ended up adopting what is basically a quadratic loss function for assessing measurement errors in their study. This is a very common loss function assumed in statistics. Similarly, Datar and Gupta ranked their estimators on the basis of two statistical criteria (i.e., lack of bias and consistency).

function assumed in statistics. Similarly, Datar and Gupta ranked their estimators on the basis of two statistical criteria (i.e., lack of bias and consistency).

Nevertheless, efforts to identify more specific decision criteria to apply in empirical research in managerial accounting should not cease. Such efforts may be possible, for example, in constrained decision settings wherein more economically relevant benchmarks can actually be applied in evaluations of alternative cost-accounting numbers. Whatever criteria are adopted, the main challenge facing researchers in managerial accounting in the future will be to identify substantive issues of a general nature that can be addressed analytically and empirically, thereby establishing both practical and theoretical implications of their results.

References

Babad, Y. M., and B. V. Balachandran. 1993. Cost driver optimization in activity-based costing. *The Accounting Review* 68 (July): 563–75.

Banker, R. D., and H. H. Johnston. 1993. An empirical study of cost drivers in the U.S. airline industry. *The Accounting Review* 68 (July): 576–601.

Datar, S. M., S. Kekre, T. Mukhopadhyay, and K. Srinivasan. 1993. Simultaneous estimates of cost drivers. *The Accounting Review* 68 (July): 602–14.

——, and M. Gupta. 1992. Aggregation, specification and measurement errors in product costing. Working paper, Washington University, St. Louis, MO.

Demski, J., and J. Christensen. 1992. The classical foundation of "modern" costing. Working paper, Yale University, New Haven, CT.

Dopuch, N., J. Birnberg, and J. Demski. 1967. An extension of standard cost variance analysis. *The Accounting Review* 42 (July): 526–36.

Hwang, Y., J. H. Evans III, and V. G. Hegde. 1992. Product cost bias and selection of an allocation base. Working paper, University of Pittsburgh, PA.

Vives, X. 1990. Trade association disclosure rules, incentive to share information, and welfare. *Rand Journal of Economics* 21 (Autumn): 409–30.

SHORTER ARTICLES

1. Ray Ball, S. P. Kothari, and Ross L. Watts
   Economic Determinants of the Relation Between Earnings Changes and Stock Returns

2. Mimi L. Alciatore
   New Evidence on SFAS No. 69 and the Components of the Change in Reserve Value

3. Wayne H. Shaw and Heather A. Wier
   Organizational Form Choice and the Valuation of Oil and Gas Producers

4. Lawrence D. Brown and Kwon-Jung Kim
   The Association Between Nonearnings Disclosures by Small Firms and Positive Abnormal Returns

5. Mario J. Maletta and Thomas Kida
   The Effect of Risk Factors on Auditors' Configural Information Processing

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 622–638

Economic Determinants of the Relation Between Earnings Changes and Stock Returns

Ray Ball
S. P. Kothari
Ross L. Watts
University of Rochester

SYNOPSIS AND INTRODUCTION: In competitive product markets, product prices and thus firms' revenues incorporate the cost of equity capital. In a competitive capital market, the cost of equity capital (the expected return on equity) increases with the risk of firms' investments. Because accounting earnings are calculated without deducting the cost of equity capital, they are expected to be an increasing function of firms' investment risks. This simple competitive equilibrium analysis predicts a positive relation between changes in investment risk and expected earnings. The presence of corporate debt complicates the analysis because leverage effects seem likely to affect the relation between changes in investment risk and expected earnings.

Using annual earnings and return data from 1950 to 1988, we document a statistically significant positive association between changes in equities' relative risks and in earnings. However, on average, only a small proportion of changes in earnings can be attributed to changes in risk. A much larger proportion is attributable to changes in economic rents (windfall gains and losses). The observed positive association between changes in earnings and changes in equities' risks suggests that leverage effects do not fully offset the effect of changes in investment risks. This

We are grateful for the comments of participants in workshops at the University of California at Berkeley, University of Chicago, Columbia University, Baruch College CUNY, Duke University, Michigan State University, New York University, University of Southern California, SUNY at Buffalo, University of Washington at Seattle, and the 1988 European Finance Association meetings at Istanbul. We also benefitted from the comments of Mike Barclay, Vic Bernard, John Bildersee, Jim Brickley, Dan Collins, George Foster, Robert Holthausen, Wayne Landsman, Richard Leftwich, Wayne Marr, Jim Ohlson, Katherine Schipper, Terry Shevlin, Dee Shores, John Wild, and Jerry Zimmerman. We acknowledge excellent research assistance by David Atlas, Tony Greig, and Richard Sloan. We acknowledge financial support from the Managerial Economics Research Center at the Simon School, University of Rochester, and from the John M. Olin Foundation.

Submitted January 1991.
Accepted September 1992.

association is robust with respect to subperiod analysis, alternative specifications of the earnings change variable, alternative data-availability requirements, and the number of portfolios formed.

Key Words: *Earnings, Stock prices, Risk, Leverage, Windfall gains and losses.*

Data Availability: *All data used are publicly available from sources identified in the text.*

IN section I, we discuss the economic determinants of changes in accounting earnings and propose that they are associated with changes in equity risk. In section II, our research design, data, and sample selection process are described. The empirical results are given in section III, and the results of a range of diagnostic tests are given in section IV. Conclusions are presented in section V.

I. Economic Determinants of Earnings Changes and Their Implications

Standard valuation models (see, e.g., Fama and Miller 1972, chap. 2) suggest that earnings changes can be attributable to changes in expected returns on equity (the cost of equity capital) as well as to changes in economic rents (windfall gains and losses). However, research on the relation between changes in earnings and stock returns has implicitly interpreted changes in earnings as information about economic rents, commonly describing them as earnings "news," "surprises," or "innovations."[1] We propose that changes in expected earnings and stock returns are jointly determined by changes in firms' investment risks and in their leverage. Thus, part of the observed association between changes in earnings and stock returns is predicted to be attributable to changes in equities' risks and, hence, expected returns. We discuss the relation between risk and changes in earnings initially in the context of an all-equity capital structure, then add the complication of debt financing.

Investment Risk Changes as Determinants of Changes in Earnings

A firm's investment risk is the market-value-weighted sum of the risks of its individual investments. The weights vary with the amounts invested in new projects, with the success (and hence value) of individual projects or lines of business, and with mergers, acquisitions, divestitures, restructurings, and plant closings. In addition, investment risk is affected by changes in operating leverage (see, e.g., Christie 1989; Lev 1974), which varies with product prices and factor costs. Finally, investment risk varies over time as individual products or lines of business mature or are repositioned in their markets. In summary, for a variety of reasons we expect that firms' investment risks, and thus the expected returns on their equity, change continuously.

---

[1] See Foster (1977), Beaver et al. (1979, 1980). The random-walk earnings model in Ball and Watts (1972) encourages an economic rents interpretation of earnings changes, as does the earnings increase/decrease research design of Ball and Brown (1968). However, in discussing the form of the earnings-returns relation, Ball and Brown (1968, fn. 40) conjecture that "income numbers convey information about the convariablity of the income process." More recently, the literature on earnings response coefficients recognizes that riskiness of the earnings stream affects the price-earnings relation (see, e.g., Collins and Kothari 1989; Easton and Zmijewski 1989; Kormendi and Lipe 1987).

We also expect variations in expected returns to be reflected in accounting earnings changes because accounting earnings are calculated without deducting the cost of equity capital. In competitive product markets, product prices, and thus firms' sales revenues, reflect the cost of capital for the level of risk involved. Without any deduction of cost of capital, earnings will also reflect the cost of capital and thus risk. If a firm increases its proportion of sales and investment in a higher-risk line of business, it not only increases its relative risk, but also its expected accounting return.[2]

Since changes in expected return are reflected in both equity returns and accounting earnings, we hypothesize that changes in investment risk induce a relation between returns and changes in accounting earnings.

Effect of Leverage

In a Miller-Modigliani world without taxes, leverage does not affect the *firm's* expected return, but it does affect the *stock's* risk and expected return, which raises the potential for another form of relation between changes in earnings and stock returns. Specifying the effect of leverage on the relation between changes in earnings and returns is complex and requires general equilibrium modelling. We consider only the direct effect of leverage on the relation between returns and changes in earnings. In this simple setting, leverage affects the interest expense deducted from earnings as well as the risks of firms' equities. A negative relation between changes in earnings and in equity risk is implied. This negative effect of leverage is hypothesized to offset, at least in part, our earlier expectation of a positive relation between change in earnings and in equity risk.

II. Research Design

We require a research design that is sensitive to annual changes in risk, so we use an annual-returns version of the Ibbotson (1975) risk-estimation technique (see Ball and Kothari 1989, for further details).[3] We investigate the association between annual changes in earnings (scaled by price) and stocks' risks and returns in the year before, the year of, and the year following the changes in earnings.

Risk and Abnormal Return Estimates

All firms on COMPUSTAT with December 31 fiscal year-ends are ranked on their unexpected earnings (defined below) in each of the 37 years during the 1951–1987 period, and are assigned to decile portfolios in equal numbers. The first (tenth) portfolio therefore is rebalanced annually to contain each year's 10 percent worst (best) earnings performers. The earnings-performance year is designated as year 0 in event time and is the year whose earnings are used to sort firms into portfolios. For event year 0, the annual buy-and-hold excess return (net of the riskless rate) on each of the ten earnings-performance portfolios is calculated for each of the 37 calendar years. From this time series, we estimate the beta and average annual abnormal return for each portfolio. These statistics are, therefore, functions only of each earnings-performance portfolio's returns *in the year in which its earnings performance occurs*; returns in no other event-year are used in obtaining the estimated event-year risks and abnormal returns.

---

[2] The simple economic model of the firm implied in the above reasoning abstracts from several factors. First, it ignores the effects of imperfect accounting methods and mandatory or voluntary changes in accounting method. Second, it ignores that part of investment risk variation that is inframarginal and absorbed by the firm, rather than passed on to consumers.

[3] Annual returns are used because annual-return beta estimates appear more consistent with the CAPM and are effective in identifying risk changes in a similar context. Handa et al. (1989) show that, when returns are measured annually, there is more dispersion in estimated betas, the cross-section risk/return intercept is more consistent with CAPM predictions, and the "size effect" is insignificant. Ball and Kothari (1989) show that, for portfolios formed on the basis of ranked *ex post* security returns, annual-return betas estimated from the same procedure used here successfully identify associated annual risk (and expected return) changes. *Ex post* security return portfolios overlap *ex post* earnings change portfolios, so the latter results are particularly relevant.

The above procedure is repeated for event years −1 and +1. A time series of 37 annual excess returns is constructed for each portfolio, and market model parameters are estimated, over 1950–1986 for the event year −1, and over 1952–1988 for event year +1. The three resulting 37-year time series (for event years −1, 0, and +1) have two important properties. First, for each earnings-performance portfolio, the three time series are independent in that they utilize non-overlapping security returns. Second, individual return observations within each time series are non-overlapping, so the procedure does not induce autocorrelation in the market model regression residuals.

We estimate the market model in risk-premium form to measure relative risk, thus controlling for the effects of interest rates and the market risk premium on returns. The market model for portfolio p $(p=1,\ldots,10)$ in event-year τ $(\tau=-1,0,+1)$ is:

$$(R_{pt}-R_{ft})=\alpha_{p,\tau}+\beta_{p,\tau}(R_{mt}-R_{ft})+\epsilon_{pt}, \tag{1}$$

where R_{pt} = equal-weighted return on portfolio p in year t; R_{mt} = equal-weighted return on the market portfolio in year t; R_{ft} = riskless rate of return in year t; $\alpha_{p,\tau}$ and $\beta_{p,\tau}$ are the intercept and slope coefficients in event year τ; and ϵ_{pt} is the error term. The intercept $\alpha_{p,0}$ is the estimated abnormal return of portfolio p in the earnings-performance year $(\tau=0)$, under the maintained hypothesis of the capital asset pricing model (CAPM). The slope $\beta_{p,0}$ is portfolio p's relative risk during the earnings-performance year. Neither $\alpha_{p,0}$ nor $\beta_{p,0}$ is influenced by returns in any other event-time year. To assess the risk variation from the year before to the year after the earnings-performance year, we also estimate equation (1) in event years $\tau=-1$ and $+1$.

Although we use data over a 37-year period, we neither assume parameter stationarity for *individual* stocks nor require that each stock be available throughout the 37-year period. Instead, parameter stationarity is assumed at the portfolio level. For example, the beta of the worst-performance portfolio is assumed constant over chronological time, even though its composition of stocks varies every calendar year.[4] Stationarity is assumed only at the relation between earnings and market return parameters, the process being studied.

Sample Selection, Data, and Variable Definitions

The sample is all firms with a December 31 fiscal year-end, earnings data for at least six years during 1950–1988, and monthly return data on the Center for Research in Security Prices (CRSP) monthly tape. Earnings data are obtained from the COM-

---

[4] Even then, beta stationarity of earnings-performance portfolios is not a critical assumption. If the portfolio betas are not stationary and the riskiness of stocks assigned to portfolios is independent of the market return, then the portfolio betas estimated from the 37-year series are statistically less precise, but still unbiased and consistent. Nevertheless, in section IV we report subperiod results, betas estimated separately in good and bad market-performance years, and results of estimating betas separately in each year by using daily returns. These results are consistent with a positive association between risk changes and earnings changes.

PUSTAT annual industrial, research, and history tapes. The resulting sample consists of 28,294 firm-years, an average of 764 firms per year.

We use annual primary earnings per share excluding extraordinary items and discontinued operations, adjusted for stock splits and stock dividends. We ignore the effect of accounting method changes because they are relatively infrequent.[5] However, because of income smoothing or other reasons, the incidence of accounting method changes among the extreme good and bad earnings-performance stocks could be greater than among other stocks. The power of the test would be reduced to the extent that stocks are misclassified into the ten earnings-performance portfolios because of accounting method changes.

The scaled earnings change variable for firm i is defined as:

$$\Delta X_{it}=(X_{it}-X_{it-1})/P_{it-1}, \tag{2}$$

where X_{it} is earnings per share in year t, and P_{it-1} is share price at the close of year $t-1$.[6] The scaled earnings change then is orthogonalized with respect to the market return by estimating the following firm-specific time-series regression with all available data (minimum of five observations):

$$\Delta X_{it}=\gamma_{0i}+\gamma_{1i}(R_{mt}-R_{ft})+\eta_{it}, \tag{3}$$

where R_{mt} is the CRSP equal-weighted annual return obtained by summing monthly returns or the sample mean annual return on the available securities, and R_{ft} is the annual T-bill interest rate. This avoids any correlation between the market return and the assignment of stocks to portfolios that could induce a spurious association between changes in risk and changes in earnings. Earnings performance is defined as:

$$UX_{it}=\gamma_{0i}+\eta_{it}. \tag{4}$$

The ten earnings-performance portfolios are formed on the basis of annually ranked UX_{it}.

III. Empirical Results

Unexpected Earnings and Total Returns: Descriptive Statistics

We report the average earnings performance and total returns on each of the ten portfolios during $\tau=-1$, 0, and $+1$ in table 1. The CRSP equal-weighted return and the sample average return also are reported. By construction, UX increases monotonically with the portfolio rank. The distribution of the earnings-performance variable is leptokurtic. The worst portfolio's earnings change is -27.44 percent of its market value; that of the best is 26.44 percent. An implication of leptokurtosis is that the extreme earnings-performance portfolios are likely to exhibit relatively large risk changes.

---

[5] A virtually exhaustive sample of accounting method changes from 1969 to 1988 in Pincus and Wasley (1991, table 4) indicates that the probability that a firm's earnings are affected by either voluntary or mandatory accounting method changes is 0.084 (our calculation).

[6] Analysts' forecast errors are an alternative to the earnings variable used in this study. Since analysts' forecasts are unavailable in machine-readable form prior to the seventies and the early coverage is generally limited to large market-capitalization stocks, we begin with earnings change as a proxy for the market-unexpected earnings. The advantage of analysts' forecast errors would be that they are less likely to be serially correlated, particularly compared to the behavior of extreme earnings changes through time. As discussed below, the effect of ignoring the negative serial correlation in extreme earnings changes biases against the risk-change hypothesis.

Table 1
Earnings-Performance Portfolio Average Annual Returns in the Three Years Around the Earnings-Performance Year: 1951–1987

Earnings-Performance Rank (p)[a]	$UX_{p,0}$	$R_{p,-1}$[b] (S.D.)	$R_{p,0}$ (S.D.)	$R_{p,+1}$ (S.D.)	$R_{p,+1JM}$[c] (S.D.)	$R_{p,+AD}$ (S.D.)
1	−27.44%	14.5%	−2.1%	15.1%	11.0%	4.0%
		(31.6)	(27.5)	(29.1)	(16.9)	(23.6)
2	−3.76	16.7	3.6	13.2	7.6	4.9
		(28.3)	(22.1)	(24.0)	(12.5)	(18.1)
3	−1.48	16.5	5.94	13.8	6.7	6.5
		(21.4)	(20.0)	(21.2)	(11.4)	(16.1)
4	−0.39	18.0	8.4	13.7	5.7	7.6
		(20.8)	(19.1)	(18.9)	(10.4)	(14.8)
5	0.32	19.8	12.3	14.2	5.8	7.9
		(21.2)	(19.4)	(17.6)	(9.1)	(14.4)
6	1.01	20.2	17.6	16.9	6.6	9.5
		(21.2)	(20.1)	(19.4)	(9.2)	(16.4)
7	1.75	22.5	23.6	18.0	7.3	9.9
		(22.2)	(24.0)	(22.5)	(10.5)	(18.4)
8	2.93	21.8	25.8	18.5	9.2	8.4
		(23.3)	(24.0)	(22.4)	(11.6)	(27.8)
9	5.43	21.7	32.0	22.2	10.8	9.6
		(27.2)	(28.6)	(28.4)	(12.4)	(21.2)
10	26.44	10.6	43.3	25.3	15.9	7.5
		(30.4)	(38.1)	(37.4)	(16.6)	(27.0)
Equal-Weighted CRSP Average Return	—	17.8	16.7	16.8	9.5	6.6
		(26.4)	(26.5)	(26.5)	(13.2)	(20.6)
Equal-Weighted Sample Average Return	—	18.2	17.0	17.1	8.6	7.5
		(23.5)	(23.5)	(23.0)	(11.7)	(18.0)

[a] Portfolio 1 (10) consists of the 10 percent worst (best) earnings performers in every calendar year. The sample is all 28,294 firm-years with December 31 fiscal year-end and earnings and return data available on the COMPUSTAT annual industrial or research files and the CRSP monthly returns tape for at least six years. Earnings-performance years, designated as event year 0, cover 1951–1987. Each calendar year, firms are ranked on their earnings performance and assigned to ten earnings-performance portfolios. Earnings performance, $UX_{p,0}$, is earnings change, scaled by year-beginning price and orthogonalized with respect to the annual market return; see equations (2) to (4).

[b] Averages of the 37 annual portfolio returns over 1951–1987, for event years −1, 0, and +1 relative to the earnings-performance year. Buy-and-hold security returns inclusive of dividends are used.

[c] $R_{p,+1JM}$ ($R_{p,+1AD}$) is average return over 1951–1987 in January–March (April–December) of year +1.

As expected, portfolio returns in year $\tau=0$ increase monotonically with UX. The average return of the worst portfolio in year 0 is −2.1 percent and that of the best is 43.3 percent. The sample mean is 17 percent. Returns in year −1 exhibit an inverted-U-shaped relation with earnings rank in year 0, which suggests the presence of factors other than earnings anticipation. These could include (1) correlated variables omitted from the CAPM and (2) negative correlation between earnings changes in years −1 and 0 (Brooks and Buckmaster 1976).

Returns in year +1 include those of the earnings announcement period as well as those of the post-announcement period. We therefore expect portfolio 10's (1's) return to be highest (lowest) in year +1. As shown in table 1, returns are highest for portfolio 10, but are not lowest for portfolio 1. Table 1 also includes returns for the January to March and the April to December components of year +1, which roughly correspond to the earnings announcement and post-announcement periods. The January to March returns reveal a return seasonality for the portfolios (see, e.g., Keim 1983; Rozeff and Kinney 1976). Much of the high return in year +1 for portfolio 10 is concentrated in the January to March quarter, reflecting both the announcement and January seasonal effects. The April to December return for portfolio 1 is 4 percent, compared to 7.5 percent for portfolio 10, which might reflect either risk differences or post-announcement drift.

Risk Changes

The estimated betas and standard errors of the ten earnings-performance portfolios for years −1, 0, and +1 are reported in table 2. The evidence is consistent with the hypothesis that risk shifts occur as a function of changes in earnings.

Estimated changes in beta from year −1 to year 0 are a statistically significant function of changes in year 0 earnings. The rank correlation between $\Delta\beta_{p,0}$ and contemporaneous earnings performance is 0.98 (p-value<0.01), thus the null hypothesis that beta shifts are not associated with changes in earnings is rejected. Consistent with this pattern, the estimated betas of the first six portfolios (essentially the "bad news" portfolios) decrease, while the estimated betas of the remaining four portfolios increase. Changes in individual portfolio betas are not significant, except in the portfolios formed at the extremes of ranked changes in earnings. The estimated beta of the worst earnings-performance portfolio declines from 1.17 to 1.01; since the standard errors of β_{-1} and β_0 are 0.056 and 0.054, and the estimates are independent, the t-statistic for the −0.16 estimated change in beta is −2.05 (p-value<0.05). Similarly, the estimated beta of the best earnings-performance portfolio increases by +0.27, from 1.07 to 1.34, and the t-statistic for the difference is 2.30 (p-value<0.05). As is the case with earnings, changes in portfolio betas appear leptokurtic. Overall, these results are consistent with our hypothesis that changes in earnings and changes in risk are related.[7]

The estimated changes in year +1 beta are not significantly related to prior earnings performance. The rank correlation between $\Delta\beta_{p,+1}$ and year 0 earnings is −0.48 (p-value$=0.16$), and no individual change is significant. The negative correlation raises the possibility of a partial reversal of the beta shifts observed in the earnings-performance year. The last column of table 2 reports the overall estimated change in beta from year −1 to +1. The rank correlation between portfolio earnings and $\Delta\beta_{p,-1,+1}$ is 0.82 (p-value<0.01), which is consistent with the hypothesis that changes in earnings are positively associated with changes in risk. Portfolio 1's estimated beta decreases by

---

[7] The changes in risk ($\Delta\beta_{p,0}$s) cannot be explained by mean reversion in betas alone (see Blume 1975, Vasicek 1973, among others). Mean reversion predicts positive changes for portfolios whose year −1 beta is less than 1, and negative changes for the others. The changes in risk reported in table 2 are consistent with mean reversion in five out of ten portfolios. For two of the five instances where the sign is as predicted by mean reversion, the beta estimates in year 0 go beyond or below unity, which is inconsistent with mean reversion (mean reversion predicts that betas approach 1 from below or above).

Table 2

Relation Between Earnings Performance and Relative Risk in the Three Years Around the Earnings-Performance Year: 1951–1987

Earnings-Performance Rank (p)[a]	$UX_{p,0}$ %	$\beta_{p,-1}$[b] (Std Err)	$\beta_{p,0}$ (Std Err)	$\beta_{p,+1}$ (Std Err)	$\Delta\beta_{p,0}$[c]	$\Delta\beta_{p,+1}$	$\Delta\beta_{p,-1,+1}$
1	−27.44	1.17 (0.056)	1.01 (0.054)	1.05 (0.059)	−0.16*	+0.04	−0.12
2	−3.76	1.04 (0.050)	0.82 (0.041)	0.89 (0.033)	−0.22*	+0.07	−0.15*
3	−1.48	0.80 (0.030)	0.74 (0.043)	0.76 (0.046)	−0.06	+0.02	−0.04
4	−0.39	0.73 (0.055)	0.69 (0.046)	0.68 (0.044)	−0.03	−0.02	−0.05
5	0.32	0.74 (0.056)	0.71 (0.043)	0.62 (0.046)	−0.03	−0.10	−0.12
6	1.01	0.75 (0.051)	0.73 (0.042)	0.68 (0.049)	−0.02	−0.05	−0.07
7	1.75	0.81 (0.037)	0.88 (0.039)	0.83 (0.037)	+0.07	−0.06	0.01
8	2.93	0.84 (0.047)	0.88 (0.034)	0.82 (0.040)	+0.04	−0.07	−0.03
9	5.43	0.98 (0.056)	1.05 (0.043)	1.04 (0.048)	+0.07	−0.01	0.06
10	26.44	1.07 (0.080)	1.34 (0.085)	1.34 (0.074)	+0.27*	0.00	0.28*
Average		0.89	0.89	0.87	−0.01	−0.02	−0.02
Rank Correlations with Earnings Performance[d]					0.98	−0.48	0.82
p-Value					<0.01	0.16	<0.01

[a] See notes to table 1.

[b] $\beta_{p,-1}$ from regressing annual excess returns for portfolio p in the 37 event years −1 on the annual CRSP equal-weighted market portfolio excess return ($\beta_{p,0}$ and $\beta_{p,+1}$ are estimated similarly):

$$(R_{pt}-R_{ft}) = \alpha_{p,-1} + \beta_{p,-1}(R_{mt}-R_{ft}) + \epsilon_{pt} \qquad \text{for } p=1,\ldots,N.$$

[c] $\Delta\beta_{p,0}=\beta_{p,0}-\beta_{p,-1}$, $\Delta\beta_{p,+1}=\beta_{p,+1}-\beta_{p,0}$, and $\Delta\beta_{p,-1,+1}=\beta_{p,+1}-\beta_{p,-1}$.

[d] Spearman rank correlation between earnings performance and $\Delta\beta_{p,0}$, $\Delta\beta_{p,+1}$, and $\Delta\beta_{p,+1,-1}$; p-values that the correlations are zero assume ten degrees of freedom.

* Significant at p-value of 0.05.

−0.12 and portfolio 10's increases by 0.28. Again, only the changes in the extreme portfolios (2 and 10) are individually significant.

The beta estimates of the earnings-performance portfolios in years −2 and +2 reveal that the risk changes associated with changes in earnings are not permanent beyond year +1. This result is addressed further in section IV.

Leverage Effects

In table 3, we report average debt/equity ratios for firms in each earnings-performance portfolio, at the end of years −1, 0, and +1. Market values are used for equity,

Table 3
Earnings-Performance Portfolios' Debt-to-Equity Ratios in the Three Years Around the Earnings-Performance Year: 1951–1987

Earnings-Performance Rank (p)[a]	$UX_{p,0}$ %	$D/E_{p,-1}$[b]	$D/E_{p,0}$	$D/E_{p,+1}$	$\Delta D/E_{p,0}$[c]	$\Delta D/E_{p,+1}$
1	−27.44	1.07 (0.12)	1.30 (0.14)	1.30 (0.13)	0.23	0.00
2	−3.76	0.70 (0.07)	0.81 (0.09)	0.85 (0.09)	0.11	0.04
3	−1.48	0.65 (0.07)	0.72 (0.07)	0.74 (0.07)	0.07	0.02
4	−0.39	0.57 (0.06)	0.61 (0.06)	0.64 (0.06)	0.04	0.03
5	0.32	0.54 (0.05)	0.56 (0.05)	0.58 (0.05)	0.02	0.02
6	1.01	0.54 (0.05)	0.55 (0.05)	0.57 (0.05)	0.01	0.02
7	1.75	0.59 (0.06)	0.58 (0.05)	0.60 (0.06)	−0.01	0.02
8	2.93	0.68 (0.07)	0.66 (0.06)	0.67 (0.06)	−0.02	0.01
9	5.43	0.80 (0.07)	0.74 (0.07)	0.75 (0.07)	−0.06	−0.01
10	26.44	1.26 (0.13)	1.06 (0.10)	1.05 (0.10)	−0.20	−0.01
Rank Correlations with Earnings Performance[d]					−1.00	−0.44
p-value					<0.01	0.20

[a] See notes to table 1. The sample is reduced to 27,680 firm-year observations by requiring long-term debt and preferred stock data.

[b] $D/E_{p,-1}(D/E_{p,0}, D/E_{p,+1})$ is ratio of book value of long-term debt plus preferred stock to market value of common equity, both at the end of event year −1 (0, +1). Reported values are averages of 37 annual portfolio debt-to-equity ratios and their standard errors. Ratios exceeding 5 are set equal to 5.

[c] $\Delta D/E_{p,0} = D/E_{p,0} - D/E_{p,-1}$ and $D/E_{p,+1} = D/E_{p,+1} - D/E_{p,0}$.

[d] Spearman rank correlations between earnings performance and $\Delta D/E_{p,0}$ and $\Delta D/E_{p,+1}$; p-values that the correlations are zero assume ten degrees of freedom.

but book values are used for long-term debt and preferred stock.[8] Although none of the individual portfolio changes in leverage is significant, they exhibit a perfect negative rank correlation with earnings performance. The leverages of the first six earnings-performance portfolios increase in the earnings-performance year, while they decrease for the other four portfolios.[9] This evidence suggests that when earnings-induced changes in equity values alter firms' debt/equity ratios, the firms do not fully adjust their capital structures to previous levels. This is consistent with the results of Ball et al. (1976). Nevertheless, the significant positive correlation between changes in earnings and in equity risks (reported above) suggests that changes in leverage do not fully offset the hypothesized relation between changes in earnings and in investment risks.

---

[8] Data are for 27,680 firm-years, a reduction of 2.2 percent from the 28,294 observations used earlier, because book values of some debt and preferred stock are not available. Debt/equity ratios exceeding 5 are set equal to 5. The tenor of the results is unchanged when alternative truncation rules are used or when outliers are included.

We also examine the portfolios' leverage ratios in year +1 to determine whether firms make delayed capital structure adjustments. We do not expect a complete capital structure adjustment because there is a negative relation between investment risk and leverage (see, e.g., Smith and Watts 1992). There is no evidence in table 3 of a lagged capital structure adjustment. The absolute magnitudes of the changes in leverage are small compared to those in year 0 (none exceeds 0.05), and the rank correlation between earnings performance and change in leverage in year +1 is insignificant (-0.44, p-value$=0.20$).

Abnormal Returns

We report abnormal returns for the three years around the earnings-performance year in table 4. Not surprisingly, portfolio abnormal returns in event year 0 (denoted as $\alpha_{p,0}$) exhibit a perfect positive rank correlation with earnings. The relation is not linear. In the extreme performance portfolios, abnormal returns are similar in magnitude to scaled unexpected earnings; in the interior portfolios, they exceed scaled unexpected earnings. This is consistent with prior studies, including Beaver et al. (1979).

The pattern of the abnormal returns in the year prior to the earnings-performance year (denoted as $\alpha_{p,-1}$) is partially consistent with early anticipation by the security market, as reported by prior studies (see, e.g., Beaver et al. 1980; Collins et al. 1987; Freeman 1987). However, the abnormal returns exhibit an inverted-*U*-shaped function of changes in earnings possibly reflecting other phenomena occurring in year -1 or misspecification of the CAPM.[10] The rank correlation between portfolio abnormal returns in year -1 and change in year 0 earnings is only 0.32 (p-*value*$=0.37$).

Portfolio average abnormal returns in year +1 (denoted as $\alpha_{p,+1}$) are generally small but are consistent with post-announcement drift. Their rank correlation with change in earnings is 0.96 ($p<0.01$). These abnormal returns include pre-announcement and earnings-announcement effects in the early months of year +1, and thus are biased in

---

[9] Leverage changes are exaggerated by using book values for debt. For bad-news portfolios, market values of debt decrease because of increased bankruptcy risk, and the leverage ratios therefore are overestimated. The converse is true for good-news portfolios. The precise effect of earnings outcomes on market-valued leverage ratios is unknown. However, given the magnitude of change in the ratios of book value of debt to market value of equity for the extreme portfolios, it seems unlikely that the use of market values of debt would substantially alter the results. It is possible, though, that market-valued leverage changes are less dramatic and that correlations between leverage changes and earnings performance are weaker than those reported in table 3. A second possibility is that, since the extreme earnings-performance stocks are relatively more leveraged, they are more sensitive to macroeconomic variables, including interest rates, which potentially induce a greater correlation between their earnings performance and aggregate economic activity. A detailed examination of this issue is potentially interesting but is beyond the scope of this study.

[10] The declining abnormal returns for the good-news portfolios and the -8.07 percent abnormal return on the best-news portfolio could be explained by some firms' achieving 10th-portfolio earnings performance in year 0 by making write-offs in year -1. Write-offs usually accompany poor earnings performance at the time, which could explain the negative abnormal return in year -1, but they reduce deductions from future earnings. See Healy (1985) and DeAngelo (1988) for evidence on such write-offs. The phenomenon is commonly known as the "big bath."

Table 4
Earnings Performance and Abnormal Returns in the Three Years Around the Earnings-Performance Year: 1951–1987

Earnings-Performance Rank (p)[a]	$UX_{p,0}$	$\alpha_{p,-1}$[b] (Std Err)	$\alpha_{p,0}$ (Std Err)	$\alpha_{p,+1}$ (Std Err)	$\alpha_{p,+1JM}$[c] (Std Err)	$\alpha_{p,+1AD}$ (Std Err)
1	−27.44	−5.48** (1.65)	−18.85** (1.56)	−2.24 (1.70)	−0.61 (0.66)	−2.87 (1.18)
2	−3.76	−1.65 (1.47)	−10.91** (1.18)	−2.41 (0.96)	−1.29* (0.52)	−1.29 (0.75)
3	−1.48	1.22 (1.07)	−7.74** (1.24)	−0.22 (1.35)	−1.49 (0.55)	0.59 (1.00)
4	−0.39	3.58 (1.62)	−4.71** (1.34)	0.61 (1.28)	−1.88** (0.54)	1.83 (0.82)
5	0.32	5.32** (1.65)	−1.02 (1.24)	1.89 (1.32)	−1.00 (0.49)	2.23 (0.93)
6	1.01	5.57* (1.51)	4.04** (1.22)	3.75 (1.41)	−0.25 (0.47)	3.63** (1.05)
7	1.75	7.11** (1.09)	8.28** (1.12)	3.24 (1.08)	−0.30 (0.55)	3.71** (0.81)
8	2.93	6.05** (1.39)	10.54** (0.98)	3.79* (1.15)	0.88 (0.55)	2.33 (0.86)
9	5.43	4.14 (1.66)	14.76* (1.24)	4.98** (1.38)	2.01** (0.63)	2.99* (0.80)
10	26.44	−8.07** (2.36)	22.70** (2.47)	4.54 (2.14)	4.67** (0.95)	0.21 (1.32)
Average		1.78	1.71	1.79	0.07	1.34
Rank Correlations with Earnings Performance[d]		0.32	1.00	0.96	0.82	0.56
p-Value		0.37	<0.01	<0.01	<0.01	0.09

[a] See notes to table 1.

[b] $\alpha_{p,-1}$ is estimated by regressing annual excess returns for portfolio p in the 37 event years −1 on the annual CRSP equal-weighted market portfolio excess return ($\alpha_{p,0}$ and $\alpha_{p,+1}$ are estimated similarly):

$$(R_{pt}-R_{ft})=\alpha_{p,-1}+\beta_{p,-1}(R_{mt}-R_{ft})+\epsilon_{pt} \qquad \text{for } p=1,\ldots,N.$$

[c] $\alpha_{p,+1JM}(\alpha_{p,+1AD})$ is average abnormal return over 1951–1987 in January–March (April–December) of year +1, estimated as in footnote c above, but using only January–March (April–December) returns.

[d] Spearman rank correlations between earnings performance and $\alpha_{p,-1}$, $\alpha_{p,0}$, $\alpha_{p,+1}$, $\alpha_{p,+1JM}$, and $\alpha_{p,+1AD}$; p-values that the correlations are zero assume ten degrees of freedom.

* Significantly different from the cross-sectional average at p-value=0.05

** Significantly different from the cross-sectional average at p-value=0.01.

favor of finding a post-announcement drift. Virtually all firms announce their annual earnings by April 1, so we reestimate abnormal returns over the nine-month period from April to December of year +1 (denoted as $\alpha_{p,+AD}$) by regressing portfolio excess returns during April to December of year +1 on excess returns over the same period for the market index. The results suggest that post-announcement drift persists after controlling for risk changes, but it is weaker than reported previously. The rank correlation between abnormal returns and portfolio earnings is 0.56, with a p-value of 0.09. The weaker drift results are attributable in part to (1) a better control for risk changes, (2) assuming April earnings announcements, and (3) ignoring interim earnings announcements.

To gain some insight into the ability of risk changes to explain a portion of the post-annoucement drift observed in constant-beta research designs, consider the effect in our study of assuming risk to be constant (i.e., projecting the year -1 beta estimates into years 0 and $+1$) As shown in table 2, a portfolio that is long (short) in stocks with the 10 percent largest (smallest) changes in earnings experiences a change in risk of 0.40 over the two years [$+0.28-(-0.12)$]. Because the annual realized risk premium on the CRSP equal-weighted portfolio averaged 9.5 percent over 1951–1987, ignoring changes in risk would underestimate the annual CAPM expected return of the portfolio by 3.8 percent [$0.40 \bullet 9.5$ percent].

There is clear evidence in table 4 against the joint hypotheses that the market is efficient and that the CAPM describes expected returns. But the pattern of the estimated abnormal returns in year $+1$ does not appear consistent with *any* credible alternative hypothesis of which we are aware. We note three intriguing features of the year $+1$ abnormal returns.[11] First, $\alpha_{p,+1AD}$ is an inverted-*U*-shaped function of earnings performance. This general pattern is inconsistent with the hypothesis that the market systematically underutilizes earnings. In particular, the low abnormal returns earned by the extreme *good-news* portfolio in April to December of year $+1$ is inconsistent with prior research that suggests "drifts" in prices. The pattern begs an explanation.[12] Second, the correlation between the abnormal returns in years -1 and $+1$ is 0.72 (p-value<0.02) when using April to December abnormal returns in year $+1$ (which are unlikely to be affected by earnings announcements).[13] Third, there is a symmetric seasonal pattern in the abnormal returns of the extreme-earnings portfolios. The worst news portfolio earns only a -0.61 percent abnormal return in January to March of year $+1$, which is surprising because it includes pre-announcement and announcement effects. It earns a larger negative return of -2.87 percent in April to December. The pattern is reversed for the best news portfolio, which earns a large positive abnormal return during January to March and little during April to December. This suggests a stock return seasonality, such as the January seasonal (Rozeff and Kinney [1976], among others) or CAPM misspecification. Taken together, these three features are anomalous; that is, they escape systematic explanation.

IV. Diagnostic Tests

This section summarizes a series of diagnostic tests that examine whether the results reported in the previous section are robust.

---

[11] Year $+1$ abnormal returns are *ex post* estimates for portfolios of large numbers of stocks. Lo and MacKinlay (1990) show that misleading inferences occur when variables correlated with abnormal returns (here, ex post earnings performance) are used to construct the portfolios.

[12] Ou and Penman (1989, table 8) also use annual earnings data and report negative abnormal returns, for both extreme good- and bad-news portfolios, over a nine-month post-announcement period.

[13] Abnormal returns in years $+2$ and -2 (reported below) are not highly correlated with each other or with abnormal returns over April to December of year $+1$. The correlation between years -1 and $+1$ therefore seems unlikely to indicate a simple CAPM misspecification.

Sample Selection Issues

We obtain similar results in a sample of 360 firms with a December 31 fiscal year-end and continuous listing during 1965–1985. This increases the number of time-series observations in the regression that orthogonalizes changes in earnings with respect to the market return. We also obtain similar (but not independent) results over subperiods of 18 and 19 years, addressing the concern that the research design assumes stationarity over the 37-year period in the relation between returns and changes in earnings. We again obtain similar (but not independent) results using 20 portfolios to investigate changes in earnings that are more extreme. As expected, risk shifts are larger in the extreme portfolios. The results are virtually indistinguishable when the average return for the sample proxies for the market portfolio, thus constraining the sample average beta and abnormal return to unity and zero, repectively. Finally, we obtain similar results without orthogonalizing scaled changes in earnings with respect to annual market returns.

Analysis of Years −2 and +2

We study event years −2 and +2 to assess whether (1) abnormal returns on the earnings-performance portfolios are correlated through time and (2) risk shifts are transitory or continue beyond year +1. The estimation technique is as described in section II, with 35 instead of 37 years now available.

The portfolio abnormal returns in years +2 and −2 are not highly correlated with each other or with abnormal returns over April to December of year +1. If the abnormal returns were correlated, then CAPM misspecification would be a more likely explanation of their "drift."

The earnings-performance portfolios' systematic risk estimates in years −2, 0, and +2 are reported in table 5. Changes in risk associated with changes in earnings do not appear to be permanent beyond year +1. The worst earnings-performance portfolio's systematic risk increases from 1.11 in year −2 to 1.18 in year +2, whereas the best news portfolio's systematic risk remains unchanged at 1.12 from year −2 to +2. In conformance with our hypothesis, risk changes from year −2 to 0, $\Delta\beta_{p,-2,0}$, are positively correlated with earnings performance (rank correlation, 0.73; p-value, 0.02). These risk changes are reversed over the following two years because $\Delta\beta_{p,0,+2}$ is negatively correlated with earnings performance. Risk changes from year −2 to +2, $\Delta\beta_{p,-2,+2}$, thus exhibit no relation with earnings performance (rank correlation, 0.03, p-value, 0.93). These results suggest that firms, particularly those with extreme earnings performance, adjust their investment policies as a function of risk changes.[14] An alternative possibility is that risk changes are a short-term function of earnings performance.

Beta Estimation Issues

The tenor of the results is unchanged when we estimate portfolio betas from daily returns in each year 1965–1985.[15] The beta change magnitudes generally are lower,

---

[14] We also examine leverage in years −2 and +2. There is little evidence of a systematic adjustment of leverage beyond year +1. Leverage ratios in year −2 are very similar to those in year −1, and the leverage ratios in year +2 are similar to those in year +1. Leverage changes from year −2 to +2 are highly correlated with earnings performance (rank correlation, −0.98; p-value, 0.01), which is similar to the behavior of leverage changes from year −1 to 0 reported in table 3.

[15] CRSP daily returns are not available for the pre-1963 period.

Table 5
Relation Between Earnings Performance and Relative Systematic Risk in the Earnings-Performance Year and in Years −2 and +2 Relative to the Earnings-Performance Year: 1951–1987

Earnings-Performance Rank (p)[a]	$\beta_{p,-2}$[b] (Std Err)	$\beta_{p,0}$ (Std Err)	$\beta_{p,+2}$ (Std Err)	$\Delta\beta_{p,-2,0}$[c]	$\Delta\beta_{p,0,+2}$	$\Delta\beta_{p,-2,+2}$
1	1.11 (0.060)	1.02 (0.058)	1.18 (0.074)	−0.09	+0.16*	0.07
2	0.92 (0.046)	0.79 (0.043)	0.84 (0.046)	−0.13*	+0.05	−0.08
3	0.83 (0.058)	0.74 (0.043)	0.60 (0.055)	−0.10	−0.14*	−0.23*
4	0.76 (0.035)	0.67 (0.049)	0.65 (0.062)	+0.09	−0.01	−0.11
5	0.67 (0.041)	0.69 (0.046)	0.68 (0.055)	+0.02	−0.01	−0.01
6	0.72 (0.048)	0.69 (0.046)	0.70 (0.040)	−0.02	+0.01	−0.01
7	0.73 (0.050)	0.87 (0.033)	0.80 (0.043)	+0.14*	−0.07	+0.07
8	0.96 (0.045)	0.87 (0.052)	0.89 (0.043)	−0.08	+0.01	−0.07
9	1.04 (0.048)	1.00 (0.046)	0.94 (0.039)	−0.04	−0.06	−0.10
10	1.12 (0.068)	1.39 (0.093)	1.12 (0.056)	+0.26*	−0.27*	−0.01
Average	0.89	0.87	0.84	−0.01	−0.03	−0.05
Rank Correlation with Earnings Performance[d]				0.73	−0.55	0.03
p-Value				0.02	0.10	0.93

[a] See notes to table 1.

[b] $\beta_{p,-2}$ is estimated by regressing annual excess returns for portfolio p in the 37 event years −1 on the annual CRSP equal-weighted market portfolio excess return ($\beta_{p,0}$ and $\beta_{p,+2}$ are estimated similarly):

$$(R_{pt}-R_{ft})=\alpha_{p,-2}+\beta_{p,-2}(R_{mt}-R_{ft})+\epsilon_{pt} \qquad \text{for } p=1,\ldots,N.$$

[c] $\Delta\beta_{p,-2,0}=\beta_{p,0}-\beta_{p,-2}$, $\Delta B_{p,0+2}=\beta_{p,+2}-\beta_{p,0}$, and $\Delta\beta_{p,-2,+2}=\beta_{p,+2}-\beta_{p,-2}$.

[d] Spearman rank correlations between earnings performance and $\Delta\beta_{p,-2,0}$, $\Delta\beta_{p,0,+2}$, and $\Delta\beta_{p,+2,-2}$; p-values that the correlations are zero assume ten degrees of freedom.

* Significant at $\alpha=5$ percent.

which is expected because betas estimated from daily returns are biased toward unity by the larger relative effect of market microstructure on measured returns (see, e.g., Ball 1977; Foster et al. 1984).

Our earnings-performance variable, defined in equation (4), is orthogonalized with respect to the market return, in part to mitigate concern that our portfolio formation procedure induces a spurious association between changes in risk and changes in earnings. A dependence between changes in risk and earnings might arise if the stocks assigned to good-news portfolios are relatively high-risk in years when the market performs well and low-risk when it performs poorly, and conversely for bad-news portfolios. With bivariate normality, portfolio beta estimates are unbiased even if there is a dependence between the market return and the beta of stocks assigned to a portfolio. However, higher-order dependence between the market performance and portfolios' risks could yield biased beta estimates.[16] We therefore perform several diagnostic tests.

First, to verify the orthogonalization procedure at the portfolio level, we estimate correlations between each portfolio's earnings performance and the market return, using the time series of 37 years. *None* of the ten portfolios' earnings performance is reliably correlated with the market through time. The estimated correlation for the extreme good-news portfolio is only 0.11 (p-value, 0.50), and it is negative for the extreme bad-news portfolio. Thus, *a priori*, there is little concern that the estimated betas are biased.

Second, we reestimate the event-year 0 betas of each portfolio, allowing them to differ with the sign of the realized market risk premium. If the worst (best) earnings-performance portfolio consists of more (less) risky stocks in years with low market returns, then its beta will be higher (lower) in such years. However, *both* the worst and the best earnings-performance portfolios' estimated betas are lower in years with low market returns, which is inconsistent with the bias argument. The rank correlation between earnings performance and the deviation of the portfolios' betas from their 37-year average is only -0.21 (p-value, 0.56) in the years with negative risk premiums.[17] Further, the rank correlation between earnings performance and the *change* from year -1 to 0 in the deviations from their average is 0.12 (p-value, 0.75). The corresponding correlation coefficient when using the change from year 0 to $+1$ is 0.35 (p-value, 0.33); from year -1 to $+1$, it is 0.47 (p-value, 0.17).

Third, we estimate the time-series correlation between the annual market return and the calendar-year beta of each earnings-performance portfolio, which is estimated with daily or monthly returns within each calendar year. The time series is 25 years for daily betas and 37 years for monthly betas. For both daily and monthly returns, the point estimates of the correlation between betas and market returns are positive for both the worst- and best-performance portfolios. This is similar to the finding reported previously that annual beta estimates of both portfolio 1 and 10 are lower in years of poor market returns. The correlation between the market return and betas estimated from daily returns is insignificant for both portfolio 1 and 10. The beta of the worst-performance portfolio, when estimated from daily returns, correlates reliably and *positively* with the annual market return. None of the remaining nine betas exhibit a reliable correlation with the market return.

Overall, a variety of diagnostic tests yields no reliable evidence that the research design is biased in favor of finding risk changes consistent with our hypothesis.

V. Conclusions

Changes in earnings have systematic economic determinants that are likely to be associated with variation in securities' expected returns, particularly since earnings is the accounting return on equity. Identifying the economic determinants of earnings variation should improve our understanding of the earnings-price relation.

---

[16] We thank an anonymous referee for raising this issue.

[17] By construction, the corresponding correlation in the positive risk-premium years is identical in magnitude, but of opposite sign.

Ignoring endogenous variation in expected returns implies the assumption that earnings variation is attributable entirely to variation in firms' economic rents. By and large, this assumption is implicit in the "event-study" literature to date and can contribute to anomalous results. Research designs investigating cross-sectional or cross-temporal variation in earnings, or similar variables, are likely to induce variation in securities' expected returns.

This study implements a research design that allows the economic determinants of earnings to vary with the same periodicity as earnings. Specifically, we estimate and control for annual, event-time betas. We rely on the single-period CAPM, even though it is implemented over a 37-year period. The continued use of the single-period CAPM in the literature presumably reflects the absence of an empirically tractable multiperiod model.[18] A second limitation is the loss of information from data aggregation into portfolios. However, using portfolios enables more powerul tests, by allowing us to estimate annual event time betas with annual returns. A third limitation arises from errors in measuring the news in earnings. Brooks and Buckmaster (1976) show that extreme changes in earnings exhibit negative serial correlation. By ignoring this, we could assign some stocks to extreme portfolios, whereas their changes in earnings are expected to be largely reversed. We also could be failing to capture risk changes that occur earlier in event time. These measurement problems seem likely to bias the results against the risk-change hypothesis.

The results suggest a predictable positive association between changes in earnings and changes in equity risk. The results also suggest that leverage-change effects do not fully offset the hypothesized relation between changes in earnings and investment risk changes. The conclusions are robust with respect to a variety of changes in the research design.

---

[18] An anonymous referee and the editor pointed out this as a limitation of our research design. A recent study of changing-composition portfolios over long time periods is Fama and French (1992).

References

Ball, R. 1977. A note on errors in variables and estimates of systematic risk. *Australian Journal of Management* 2: 79–84.

——, and P. Brown. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6 (Autumn): 159–78.

——, and S. P. Kothari. 1989. Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns. *Journal of Financial Economics* 25 (November): 51–74.

——, B. Lev, and R. Watts. 1976. Income variation and balance sheet compositions. *Journal of Accounting Research* 14 (Spring): 1–9.

——, and R. Watts. 1972. Some time-series properties of accounting income. *Journal of Finance* 27 (July): 663–82.

Beaver, W., R. Clarke, and W. Wright. 1979. The association between unsystematic security returns and the magnitude of earnings forecast errors. *Journal of Accounting Research* 17 (Autumn): 316–40.

——, R. Lambert, and D. Morse. 1980. The information content of security prices. *Journal of Accounting & Economics* 2 (March): 3–28.

Blume, M. 1975. Betas and their regression tendencies. *Journal of Finance* 3 (July): 785–95.

Brooks, L., and D. Buckmaster. 1976. Further evidence on the time-series properties of accounting income. *Journal of Finance* 31 (December): 1359–73.

Christie, A. 1989. Equity risk, the opportunity set, production costs and debt. Working paper, University of Rochester, NY.

Collins, D., and S. P. Kothari. 1989. An analysis of intertemporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting & Economics* 11 (July): 143–81.

——, ——, and J. Rayburn. 1987. Firm size and the information content of prices with respect to earnings. *Journal of Accounting & Economics* 9 (July): 111–38.

DeAngelo, L. 1988. Managerial competition, information costs, and corporate governance: The use of accounting performance measures in proxy contests. *Journal of Accounting & Economics* 10 (January): 3–36.

Easton, P., and M. Zmijewski. 1989. Cross-sectional variation in the stock market response to the announcement of accounting earnings. *Journal of Accounting & Economics* 11 (July): 117–41.

Fama, E., and K. French. 1992. Common risk factors in the returns on bonds and stocks. *Journal of Financial Economics* (forthcoming).

——, and M. Miller. 1972. *The Theory of Finance*. Hinsdale, IL: Dryden.

Foster, G. 1977. Quarterly accounting data: Time-series properties and predictive ability results. *The Accounting Review* 52 (January): 1–21.

——, C. Olsen, and T. Shevlin. 1984. Earnings releases, anomalies and the behavior of security returns. *The Accounting Review* 59 (October): 574–603.

Freeman, R. 1987. The association between accounting earnings and security returns for large and small firms. *Journal of Accounting & Economics* 9 (July): 195–228.

Handa, P., S. P. Kothari, and C. Wasley. 1989. The relation between the return interval and betas: Implications for the size effect. *Journal of Financial Economics* 23 (June): 79–100.

Healy, R. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting & Economics* 7 (April): 85–107.

Ibbotson, R. 1975. Price performance of common stock new issues. *Journal of Financial Economics* 2 (September): 235–72.

Keim, D. 1983. Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics* 12 (June): 13–32.

Kormendi, R., and R. Lipe. 1987. Earnings innovations, earnings persistence and stock returns. *Journal of Business* 60 (July): 323–45.

Lev, B. 1974. On the association between operating leverage and risk. *Journal of Financial and Quantitative Analysis* 9 (December): 627–41.

Lo, A., and C. MacKinlay. 1990. Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies* (3): 431–67.

Ou, J., and S. Penman. 1989. Accounting measurement, P/E ratios and the information content of security prices. *Journal of Accounting Research* 27 (Supplement): 111–44.

Pincus, M., and C. Wasley. 1991. Time-series and cross-sectional characteristics of accounting changes. Working paper, Washington University, St. Louis.

Rozeff, M., and W. Kinney. 1976. Capital market seasonality: The case of stock returns. *Journal of Financial Economics* 3 (October): 379–402.

Smith, C., and R. Watts. 1992. The investment opportunity set and corporate financing, dividend, and compensation policies. *Journal of Financial Economics* 32 (December): 263–92.

Vasicek, O. 1973. A note on using cross-sectional information in Bayesian estimation of security betas. *Journal of Finance* 28 (December): 1233–39.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 639–656

New Evidence on SFAS No. 69 and the Components of the Change in Reserve Value

Mimi L. Alciatore
Southern Methodist University

SYNOPSIS AND INTRODUCTION: The purpose of this article is to investigate the information content of the change in the standardized measure (CSM) disclosure reported by oil and gas producers as assessed by market participants. Previous research by Doran et al. (1988) suggests that selected components of the CSM disclosure have a modest influence on share prices. This study reexamines this issue with a more homogeneous sample of oil and gas producers (e.g., excluding Canadian firms), and investigates all ten components of the CSM that firms must currently report. It provides evidence on the usefulness of one of the major supplemental disclosures for oil and gas producers. In particular, it shows that disaggregating the change in reserve value into the required components provides additional information over reporting only the CSM, and that six components provide incremental information relative to the other components and to net income: (1) production; (2) discoveries; (3) purchases of reserves; (4) quantity revisions; (5) price changes; and (6) the change in income taxes.

In the early 1980s, the Financial Accounting Standards Board (FASB) and the Securities Exchange Commission (SEC) decided that, although reserve value accounting data would not be reported in the financial statements, supplemental disclosures of reserve values and changes therein (i.e., the CSM disclosure) would be required because such information is not reflected in historical cost financial statements (see FASB 1982; SEC 1982). Thus, this study also contributes to the more general area regarding the incremental information in footnotes relative to information such as net income reported in the financial statements (e.g., see Beaver et al. 1982; Imhoff et al. 1992).

I gratefully acknowledge the helpful comments of M. Edgar Barrett, Randolph P. Beatty, Edward B. Deakin, James W. Deitrick, Philip Drake, Kenneth R. Ferris, John D. Martin, S. Lynne Stokes, Michael Vetsuypens, Patrick J. Wilkie, the editor, and the referees. I am also grateful for financial support from the Deloitte & Touche Foundation and the Edwin L. Cox School of Business at Southern Methodist University.

Submitted May 1991.
Accepted December 1992.

Doran et al. (1988) tested for the incremental information content, relative to historical cost earnings, of three reserve value-based measures: (1) the total annual change in reserve value (less reserve purchases and sales), as well as the change in reserve value attributable to (2) discoveries and (3) revisions in price and quantity estimates. They calculated the variables using data filed under *Statement of Financial Accounting Standards No. 69* (SFAS No. 69, FASB 1982) for 1982–1984 and found that only the revisions variable provided information on producers and that earnings did not. This study replicates their tests using SFAS No. 69 data over the same period with consistent results. The analyses of Doran et al. are then expanded by investigating the information content of the ten components of the change in reserve value required by SFAS No. 69. The set of information items currently required are more disaggregated than the set that was required prior to SFAS No. 69.[1] The study results offer new insights and provide the first evidence that gross profits from production are positively associated with returns. In addition, the findings suggest that discoveries and purchases of reserves are negatively associated with returns during the test period. Although this appears to be counterintuitive, it is consistent with analyses of the oil and gas industry by McConnell and Muscarella (1985), Picchi (1985), and Jensen (1986a, 1986b) that the market reacted negatively to oil firms investing in exploration and development (E & D) during this time period. They describe the period as one in which oil prices were declining, and firms were spending too much on E & D, and had excessive levels of reserves. Thus, additions to reserves as measured by CSM components such as discoveries need not be positively associated with returns. In addition, there is some evidence that the tax change component reflects a tax adjustment to the values of the other components.

The analysis of the CSM components is replicated using abnormal returns estimated from a market model that includes industry and market indexes. The results show that net income is positively related to abnormal returns, which indicates that earnings do provide information about firm-specific events for producers. The results also indicate that the six CSM components described above are associated with the dual-index returns.

Key Words: *Oil and gas, Standardized measure, Reserve recognition accounting, Information content.*

Data Availability: *All data must be obtained from the sources identified in the text.*

---

[1] Doran et al. (1988) also calculated their variables using reserve recognition account (RRA) data for 1979–1981. Oil and gas firms were required to report RRA reserve value data for fiscal years ending after December 25, 1979 through years ending before December 15, 1982, when SFAS No. 69 became effective. It is not relevant to examine RRA-based data because RRA rules are no longer in effect. The measures Doran et al. examined relate more to the RRA disclosures than to the data items currently provided under SFAS No. 69. In general, the CSM components involve a lower level of aggregation than the RRA components of reserve value changes (e.g., under RRA, quantity revisions were often combined with other items). In addition, some CSM components (e.g., income taxes and changes in future development costs) were not reported under RRA rules. See table 1 and appendix B for a description of the CSM components required under SFAS No. 69; see Accounting Series Release No. 269 (SEC 1979) for the required RRA disclosures.

THE remainder of this article is organized as follows. The motivation for the study is discussed in section I, and the information content of the ten components of the *CSM* is examined in section II. The last section provides a discussion and summary of the findings.

I. Background and Motivation

Previous research by Doran et al. (1988) focused on the incremental information content of three reserve value-based measures relative to historical cost earnings. Their motivation was to test assertions by the SEC and the FASB that disclosures of reserve value and changes therein required by SFAS No. 69 (and formerly by RRA) provide relevant information that is not reflected in historical cost financial statements.[2] Doran et al. conjectured that each of the three variables they examined would be positively associated with abnormal returns to the extent that the variable reliably measures unexpected changes in future cash flows resulting from the given change in reserve value (e.g., from discoveries). Several other studies (Jensen 1986a, 1986b; McConnell and Muscarella 1985; Picchi 1985) assert that oil firms were, at the margin, undertaking E & D projects with negative net present values. McConnell and Muscarella (1985) found that announcements of increases (decreases) in budgeted capital expenditures for E & D during 1975–1981 were associated with negative (positive) excess stock returns.[3] They suggest that the oil firms in their sample were overinvesting in E & D and that, consequently, the market reacted favorably (unfavorably) when the firms decreased (increased) their E & D expenditures. Picchi (1985) examined rates of return on E & D expenditures for 30 large oil companies during 1982–1984 and reported that the average ratio of the *CSM* discoveries component (*DISCOV*) to E & D expenditures ranged from approximately 0.6 to 0.9. That is, the industry was realizing, on average, only 60–90 cents on every dollar invested in E & D. On the basis of his analysis, Picchi (1985, 6) concluded, "For 23 of the [30] companies in our survey, we would recommend immediate cuts of perhaps 25–30% in exploration and production spending. It is clear that much of the money that these firms spent last year on petroleum exploration and development yielded subpar financial returns—even at $30 per barrel, let alone today's $26–$27 per barrel price structure." In addition, Magliolo (1986, 96) found that, for 1979–1983, the association between *DISCOV* and changes in the market value of oil and gas properties was significant but in the opposite direction from that predicted by his theoretical model. Finally, Jensen (1986a, 326–27) describes the economics of the oil industry during the early eighties as follows: "Crude oil reserves (the industry's major asset) were too high, and cutbacks in . . . E & D expenditures were required. . . . Instead, the industry continued to spend heavily on E & D activity even though average returns were below the cost of capital."[4] In sum, these authors describe the early eighties as a period of declining oil prices, and overspending on E & D such that oil firms had excessive levels of reserves. This view of the industry explains why additions

---

[2] See Alciatore (1990) for a review of research investigating the reliability and relevance of reserve value accounting data. Also see Spear (1992).

[3] In contrast, for all the other industries examined, McConnell and Muscarella (1985) found that increases in capital expenditures were associated with positive returns.

[4] Jensen's line of analysis is similar to that presented by Barrett (1989) and the Harvard Business School (1984) in case studies of Gulf Oil Corporation.

additions to reserves during this period, as measured by the CSM components (e.g., *DISCOV*) may not be positively associated with market returns.

II. This Information Content of the SFAS No. 69 *CSM* Components

SFAS No. 69 requires oil and gas (O & G) producers to report two supplemental reserve value-based disclosures in their annual reports: (1) an estimate of the value of their proved reserves at fiscal year-end, called the "standardized measure" (SM),[5] and (2) the "change in the standardized measure" (CSM) disclosure, which presents the total change in the firm's reserve value (i.e., in the SM) during the fiscal year, as well as the dollar effect of each significant factor (i.e., CSM component) contributing to the change. SFAS No. 69 requires that ten CSM components be separately presented.

The following pooled cross-sectional regression models were estimated for the period 1982–1984 to test for the information content of the CSM and its components:

$$CAR_{it}=\alpha+\alpha_{83}+\alpha_{84}+\beta_1 CSM_{it}+\beta_2 dNI_{it}+e_{it}, \quad (1)$$

$$\begin{aligned}CAR_{it}=\alpha+\alpha_{83}+\alpha_{84}&+\beta_1 PROD_{it}+\beta_2 DEV_{it}+\beta_3 DISCOV_{it}+\beta_4 PURCH_{it}\\&+\beta_5 SALES_{it}+\beta_6 QTYREV_{it}+\beta_7 PRICE_{it}+\beta_8 DEVCHG_{it}\\&+\beta_9 TAXCHG_{it}+\beta_{10} DISCT_{it}+\beta_{11} OTHER_{it}+\beta_{12} dNI_{it}+e_{it},\end{aligned} \quad (2)$$

where CAR_{it} is the cumulative abnormal return for firm *i* in year *t*, α_{83} and α_{84} are dummy intercepts that are used to control for year effects, the *CSM* components and the other independent variables are as defined in table 1, and e_{it} is the error term. *CAR*s were calculated by accumulating weekly prediction errors from the market model over the fiscal year.[6] The CRSP daily returns file provided the stock return data and the remaining data were obtained from firms' annual reports and Form 10Ks.[7]

---

[5] The measure is "standardized" in that firms must follow a common set of assumptions when calculating it. For instance, SFAS No. 69 requires the use of current prices and costs (as of the current balance sheet date) in estimating net cash flows from future O & G production, and the cash flows are to be discounted to the current balance sheet using a 10 percent rate. The SM is similar in construction to the RRA reserve value estimate. The main difference is that the SM is net of expected income taxes whereas the RRA estimate was on a pre-tax basis. The SEC's (1982) current reporting rules for O & G producing firms require compliance with SFAS No. 69.

[6] The independent variables were deflated by the market value of equity at the beginning of the fiscal year to control for differences in firm size that could result in heteroscedasticity (see Christie 1987). As in Doran et al. (1988), *dNI* is the deflated change in net income: (current *NI* minus previous *NI*)/market value of equity at the beginning of the current year *t*. Separate year intercepts are included to be consistent with Doran et al., who found the 1983 intercept to be significant. The market model parameters were estimated for each firm using weekly data from the 50 weeks preceding the test year and the CRSP value-weighted index. The *CSM* components listed in table 1 are as defined in SFAS No. 69 (par. 33), with one exception. Purchases and sales of minerals in place is one component in SFAS No. 69, but is divided into two components in this study, purchases (*PURCH*) and sales (*SALES*). One reason for distinguishing between purchases and sales is that firms frequently report them separately (though it is not required). In addition, the relation between returns and purchases may not be the reverse of returns and sales. That is, buying reserves is not necessarily the opposite of liquidating reserves (versus holding reserves). Some of the subtle technical issues involved in calculating and understanding the *CSM* components are discussed in appendix B.

[7] For a given test year, a firm was included in the sample if it (1) was primarily involved in O & G producing activities, defined as being classified under SIC code 1311 (crude petroleum and natural gas) on the CRSP tapes and being non-integrated (i.e., a majority of its identifiable assets, operating revenues, or operating profits related to its production segment); (2) had complete returns and accounting data available; and (3) was a U.S. corporation with a calendar year-end (see appendix A). As in Doran et al. (1988), the December 31 fiscal year-end requirement was imposed to control for cross-temporal differences in market and industry factors. It also allows for a more precise definition of the test period (in calendar time), and thereby facilitates interpretation of the results in the context of the economics of the period. After applying the criteria above, the final sample consisted of 82 observations for 34 firms (from a population of 164 firms): 23 firms for 1982, 32 for 1983, and 27 for 1984.

Table 1
Specification of Variables in Models 1 and 2

Variables	*Coefficient's Expected Sign*
Dependent Variable	
CAR = annual abnormal stock return, measured over the fiscal year	
Independent Variables	
dNI = change in net income from the previous fiscal year	+
CSM = change in the standardized measure of discounted future net cash flows relating to proved oil and gas reserves	+or−
Components of the CSM	
PROD = sales and transfers of oil and gas produced during the period, net of production costs	−
DEV = previously estimated development costs incurred during the period	−
DISCOV = net change attributable to extensions, discoveries, and improved recovery	+
PURCH = net change attributable to purchases of minerals in place	+
SALES = net change attributable to sales of minerals in place	+
QTYREV = net change attributable to revisions in (reserve) quantity estimates	+
PRICE = net change in sales and transfer prices and in production (lifting) costs related to future production	+
DEVCHG = changes in estimated future development costs	+
TAXCHG = net change in income taxes	+or−
DISCT = accretion of discount	0
OTHER = other—unspecified	0

Note: Each *CSM* component measures the change in the standardized measure that is attributable to that factor. For instance, the *DEVCHG* component measures the amount of the *CSM* resulting from changes in estimated future development costs. The *CSM* could have either a postive or negative coefficient because it is the sum of the components, some of which have negative expected signs and some of which have positive ones. Similarly, *TAXCHG* is a tax adjustment on the other components. For a given component, the tax adjustment would be expected to have the opposite association with returns as the component. Thus, *TAXCHG* could have either a positive or negative coefficient because it is the sum of a set of opposing effects to the other coefficients, which have negative and positive expected signs.

Table 2 presents the regression results for models 1 and 2. In general, the results present a different picture of the CSM disclosure from that suggested by previous research. First, I tested for the incremental information content of the *CSM* components as a group plus *dNI* (change in net income) relative to the *CSM* and *dNI* (see Lipe 1986). The *F*-statistic for incremental information content comparing model 2 with model 1 was 2.06 (0.04 significance level), which suggests that disaggregating the CSM into the required components provides additional information over reporting only the total change in reserve value (see section III for further discussion).[8] Table 2 also

---

[8] This finding is robust across numerous replications including reestimation with: (1) weighted least squares (WLS) and (2) ordinary least squares (OLS) for a subsample comprised of the observations from 1983–1984 (a more homogeneous sample than 1982–1984). However, note that the test is essentially a test for

Table 2
OLS Regression Results for Models 1 and 2

	α	α_{83}	α_{84}	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}
Model 1:[a]															
Coefficient	−0.36	0.94	−0.04	−0.03	0.56										
White's Probability	0.01	<0.001	0.81	0.90	0.11										
R-square	0.48														
Adjusted *R*-square	0.45														
F-statistic	17.62														
Probability	<0.001														
Model 2:[b]															
Coefficient	−0.53	0.74	−0.12	−1.50	−2.01	−0.94	−1.27	−0.18	−1.00	−0.99	0.49	−1.66	0.39	−0.01	0.36
White's Probability	0.002	<0.001	0.42	0.03	0.16	0.02	0.01	0.69	0.01	0.01	0.52	0.002	0.49	0.99	0.12
R-square	0.60														
Adjusted *R*-square	0.52														
F-statistic	7.20														
Probability	<0.001														

Model 1 vs. Model 2 (*F*-test for incremental explanatory power) $F = 2.06$ (0.04 significance level)

Note: The dependent and independent variables are defined in table 1. The total number of observations is 82 (23 in 1982, 32 in 1983, and 27 in 1984). The probability values are based on two-tailed *t*-tests except those for the intercepts (one-tailed).

[a] $CAR_{it} = \alpha + \alpha_{83} + \alpha_{84} + \beta_1 CSM_{it} + \beta_2 dNI_{it} + e_{it}$.

[b] $CAR_{it} = \alpha + \alpha_{83} + \alpha_{84} + \beta_1 PROD_{it} + \beta_2 DEV_{it} + \beta_3 DISCOV_{it} + \beta_4 PURCH_{it} + \beta_5 SALES_{it} + \beta_6 QTYREV_{it} + \beta_7 PRICE_{it} + \beta_8 DEVCHG_{it} + \beta_9 TAXCHG_{it} + \beta_{10} DISCT_{it} + \beta_{11} OTHER_{it} + \beta_{12} dNI_{it} + e_{it}$.

shows that the *CSM* and *dNI* are statistically insignificant, which is consistent with the findings of Doran et al. (1988). In addition, Spear (1992), using an event study methodology, also found that the *CSM* did not provide information at the time of its release.

Each *CSM* component was examined to test for incremental information content relative to the other components and *dNI*. Table 2 shows that six of the components are significantly negatively related to abnormal returns: *PROD*, *DISCOV*, *PURCH*, *TAXCHG*, *QTYREV*, and *PRICE*.[9] *PROD* and *DEV* have negative coefficients, as expected, but *DEV* is insignificant.[10] The negative coefficients of several of the other significant components are consistent with the industry's economic environment during the period. Spot prices per barrel for the benchmark West Texas Intermediate (WTI) crude oil declined over 1982–1984 by approximately 9, 7.3, and 11 percent, respectively. For example, during 1984, the price declined from $29.60 to $26.40 per barrel, and further declines were observed in 1985–1986. Such magnitudes of price declines can have major effects on the industry; many E & D programs are undertaken on the basis of assumed expected prices and subsequent changes in prices can significantly affect the profitability of the projects (see Picchi 1985, 1–2). Thus, a likely explanation for negative associations between returns and variables such as *DISCOV* and

---

the equality of coefficients and that equality could also be rejected if the degree of measurement error varies across components. For example, this could occur if the expectation model for the components is more valid for some components than for others. Finally, although firms' costs in disaggregating the *CSM* into its components are unknown, one independent O & G producer estimates that it takes a top-level accountant approximately three days to prepare the *CSM* disclosure when using a tailored computer program, not counting the time it takes employees to gather all of the underlying data. Further, a major oil firm estimates that preparing the *CSM* disclosure requires approximately 4 to 8 man-weeks of employee time. These estimates suggest that the total cost of preparing, reporting, and auditing (i.e., reviewing) the *CSM* disclosure is often nontrivial.

[9] The Goldfeld-Quandt test for heteroscedasticity indicated that 1982 is different from 1983–1984 for models 1 and 2. When the models were reestimated with WLS and OLS for the 1983–1984 subsample, both sets of results were very similar to the original findings (and coefficients were often more significant). Potential outliers were identified by using techniques suggested by Belsley et al. (1980) and were deleted; the findings remained the same. Also, an accounting method dummy variable (successful efforts or full cost) was added to the models to control for differences related to accounting methods (see Deakin 1979), and the results remained the same. Tests of structural change indicated that the null of no change in the slopes over the years was rejected for model 1 (0.05 level) but not for model 2. Finally, regarding cross-sectional dependence in the residuals, there are not enough time-series data in this study to estimate the variance-covariance matrix without introducing a substantial amount of estimation error into the matrix estimate (see Bernard 1987). The approach of Doran et al. (1988, 396–98) is not used here because the results suggest that the assumption that the correlations in the dependent variable translate to the residuals may not hold here. The year intercepts may remove some of the potential dependence. Further, the use of a dual-index abnormal returns model (presented subsequently) is likely to eliminate a major source of cross-sectional dependence in the returns.

[10] *PROD* measures gross profits relating to production. Because production decreases the SM, *PROD* is (by definition) a negative component of the *CSM*. The more negative *PROD* is, the higher the firm's gross profit. A negative coefficient for *PROD* indicates that gross profits are *positively* associated with returns. (The coefficient could be positive if gross profits were less than expected [i.e., bad news].) Note that average gross profits increased from 1982 to 1983 (i.e., *PROD* decreased from −0.28 to −0.40); they decreased from 1983 to 1984 (to −0.36), but remained higher than in 1982. *DEV* is a positive component (it increases the SM). To the extent that the market viewed firms as spending too much on development during the period as suggested by Jensen (1986a, 1986b), *DEV* would be predicted to have a negative coefficient. *DEV* did have a significantly negative coefficient in some replications of the study (e.g., if *DEV* is substituted for *COSTS* in models 3 through 5). *DEVCHG* and *SALES* may be insignificant because of a lack of variation in these variables; they have 40 and 54 zero values, respectively (zeros, not missing values). Because the *DISCT* component has a fixed value, it is expected to be insignificant, as is *OTHER* because it is merely a plug figure. Similarly, the *CSM* is the sum of so many different components that it may represent a rather garbled information measure (see section III). Also, note that the models in this study assume that changes in non-reserve assets are insignificant and do not comprise a correlated omitted variable. The fact that the sample firms are primarily producers (e.g., vs. refiners) supports this assumption.

PURCH is that investments in reserves during the period were viewed unfavorably by the stock market, in part because of the declining trend in prices. That is, the net effects of unexpected changes in current and future cash flows relating to discoveries and purchases, including price effects, were negative. The coefficient for *DISCOV* (*PURCH*) in model 2 suggests that an increase in *DISCOV* (*PURCH*) of one standard deviation above its mean would be associated with a decrease in the abnormal return from −0.0087 to −0.1728 (to −0.2185 for *PURCH*).[11]

This interpretation is consistent with that of Jensen (1986a, 1986b) and other studies discussed earlier. For example, Picchi's results showed that the ratio of *DISCOV* to E & D expenditures was less than one, and his negative recommendation as a stock analyst is also consistent with the negative coefficients for *DISCOV* and *PURCH* obtained here. Similarly, Magliolo's (1986) finding is consistent with obtaining a negative coefficient for *DISCOV*.

A basic premise used by Jensen (1986a, 1986b) is that O & G prices are a crucial determinant of the optimal level of E & D. To further explore the effect of prices as a potential explanation of the findings, the following pooled cross-sectional models were estimated:

$$\begin{aligned} CAR_{it} = {} & \alpha + \beta_1 PROD_{it} + \beta_2 DISCOV_{it} + \beta_3 PURCH_{it} + \beta_4 QTYREV_{it} \\ & + \beta_5 PRICE_{it} + \beta_6 TAXCHG_{it} + \beta_7 COSTS_{it} + \beta_8 dNI_{it} + e_{it}, \end{aligned} \tag{3}$$

$$\begin{aligned} CAR_{it} = {} & \alpha + \beta_1 PROD_{it} + \beta_2 DISCOV_{it} + \beta_3 DISCOVPP_{it} + \beta_4 PURCH_{it} + \beta_5 PURCHPP_{it} \\ & + \beta_6 QTYREV_{it} + \beta_7 PRICE_{it} + \beta_8 TAXCHG_{it} + \beta_9 COSTS_{it} + \beta_{10} dNI_{it} + e_{it}, \end{aligned} \tag{4}$$

$$\begin{aligned} CAR_{it} = {} & \alpha + \beta_1 PROD_{it} + \beta_2 DISCOV_{it} + \beta_3 DISCOVPR_{it} + \beta_4 PURCH_{it} + \beta_5 PURCHPR_{it} \\ & + \beta_6 QTYREV_{it} + \beta_7 PRICE_{it} + \beta_8 TAXCHG_{it} + \beta_9 COSTS_{it} + \beta_{10} dNI_{it} + e_{it}, \end{aligned} \tag{5}$$

where *DISCOVPP* (*PURCHPP*) in model 4 measures the interaction between *DISCOV* (*PURCH*) and the percentage change in oil prices over the year, *PP* (based on spot prices per barrel for WTI); *DISCOVPR* (*PURCHPR*) in model 5 measures the interaction between *DISCOV* (*PURCH*) and the PRICE components; and *COSTS* (models 3 through 5) are the acquisition, exploration, and development costs incurred by the firm in that year.[12]

---

[11] Discoveries (purchases) could have a negative coefficient even though positive unexpected cash flows were announced at the time of the discovery (purchase) if the market was revising the expected cash flows from the discovered (purchased) reserves downward later during the year as a result of declining oil prices. The base level of the *CAR* of −0.0087 was calculated by multiplying each regression coefficient in model 2 by the mean for that variable. The sum of these products was then added to the intercept, plus the value of each dummy year intercept (A83 and A84) multiplied by the proportion of total observations for that year (e.g., 32/82 for 1983). The value of the return after a given variable (e.g., *DISCOV*) increased was calculated by multiplying the variable's coefficient by one standard deviation for that variable and adding this product to the original base value for the return.

[12] *PP* is a general measure of oil price changes and is the same for all firms in a given year. In contrast, *PRICE* is a firm-specific measure of the effects of both oil and gas price changes in a given year and is based upon prices pertaining to the firm's specific reserve holdings. The year-end prices for WTI crude were obtained from *The Wall Street Journal* for 1983–84 and were estimated for 1982 on the basis of prices for other crudes in *The Wall Street Journal*. Note that, if additive measures of the change in oil prices for each of the years are used to control for price changes, instead of an interactive term, the measures basically function as separate year intercept terms. In contrast, an interactive variable considers both the price change for a given year and its effects given the firm's reserves (e.g., those discovered and purchased). Costs incurred were obtained from firms' SFAS No. 69 disclosures. Separate year intercepts were not used because they could remove price change effects that the interactive terms are intended to investigate.

Table 3 presents the regression results for models 3 through 5. Model 3 is an adaptation of model 2 (with the insignficant variables being deleted and the *COSTS* variable added) and its results parallel those of model 2. *COSTS* have a negative coefficient, as expected, but it is insignificant. When the price interaction variables are added in model 4, the coefficient for *DISCOV* becomes significantly positive. This suggests that, when other factors (e.g., price changes) are controlled for, discoveries are positively associated with returns because an (unexpected) asset has been added to the firm's holdings. However, during the year the asset was added, the value of that asset was decreasing because of the decline in oil prices, and *DISCOVPP* is a proxy for this decline in value.[13] Its positive significant coefficient suggests that the decline in value is associated with a negative effect on returns. The coefficients in model 3 suggest that one standard deviation increase in *DISCOV* above its mean would increase the return by 0.99 (from −0.0017). However, a one standard deviation decrease in *DISCOVPP* below its mean (i.e., representing a larger value decline) would decrease the return by 1.22. These results suggest that the positive effects on returns of discoveries were more than offset by the effects of decreases in the value of the discoveries caused by declining prices. These findings may help to explain why *DISCOV* could have a negative sign in models 2 and 3 if it is reflecting both the addition of reserves and the simultaneous decrease in the value of the discoveries, and why *DISCOV*'s coefficient becomes positive once the price effect is (at least partially) controlled for by adding *DISCOVPP*.

In model 5, *DISCOVPR* is used as an alternative proxy for the decline in the asset value, and its positive coefficient is consistent with the findings for *DISCOVPP* in model 4. However, *DISCOV* is insignificant in model 5. The negative coefficient for *PURCH* suggests that, on average, the market did not look favorably upon reserve purchases during this period. However, the price interaction variable, *PURCHPR*, was negative on average and significant. Its negative coefficient suggests that the more firms purchased during times of relatively greater price declines (i.e., the more negative *PURCHPR*), the higher was a (partially) offsetting positive effect on returns relating to the purchases (presumably attributable to buying reserves at a cheap price). For 1982–1984, the mean for *PURCH* was 0.026, 0.069, and 0.073, respectively, and the mean for *PRICE* was −0.016, −0.240, and −0.112, respectively. Thus, it appears that firms bought more reserves in 1983–1984 when the price decreases were larger and prices were lower. Nevertheless, oil prices continued to decline throughout and after the period, which could explain the negative coefficient for *PURCH* in models 2, 3, and 5, with the timing of the purchases providing a partial lessening of this effect as suggested by *PURCHPR*'s coefficient.[14]

TAXCHG has a significantly negative coefficient in models 2 through 5. Because it is a tax adjustment to the other components (see appendix B), its coefficient may be a composite of tax adjustments to the coefficients of the other components. The results provide some support for this interpretation. For example, current-year taxes on gross

---

[13] *DISCOVPP* is negative, as it is the product of (positive) *DISCOV* and the price decline that year (*PP* was negative in each year). For example, the more a firm discovered (and presumably the more it lost from a given price decline), the more negative its *DISCOVPP*. Similarly, the greater the decline in a given year, the more negative a firm's *DISCOVPP*.

[14] See *Oil and Gas Reserve Disclosures: 1980–1983* by Arthur Andersen & Co. (1983) for a discussion of firms' buying "cheap" reserves during this period, and an analysis of the current and expected oil and gas prices at the time (see S-11 through S-15).

Table 3
OLS Regression Results for Models 3 through 5: An Examination of Interaction Effects Relating to Oil Price Changes

	α	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
Model 3:[a]											
Coefficient	−0.28	−2.00	−1.27	−1.38	−1.14	−1.25	−1.50	−0.52	0.03		
White's Probability	0.12	0.003	0.03	0.07	0.05	0.02	0.06	0.21	0.95		
R-square	0.26										
Adjusted *R*-square	0.17										
F-statistic	3.13										
Probability	0.004										
Model 4:[b]											
Coefficient	−0.30	−1.56	5.79	69.68	−0.82	6.18	−0.99	−1.13	−1.69	−0.47	0.24
White's Probability	0.06	0.01	<0.001	<0.001	0.70	0.76	0.04	0.03	0.01	0.18	0.51
R-square	0.44										
Adjusted *R*-square	0.36										
F-statistic	5.50										
Probability	<0.001										
Model 5:[c]											
Coefficient	−0.31	−1.64	−0.59	3.93	−1.98	−1.46	−0.89	−2.07	−1.54	−0.39	−0.39
White's Probability	0.08	0.02	0.39	0.05	0.01	0.05	0.08	0.002	0.03	0.32	0.23
R-square	0.29										
Adjusted *R*-square	0.19										
F-statistic	2.87										
Probability	0.005										

Note: The dependent and independent variables are defined in table 1 and the text. The total number of observations is 82 (23 in 1982, 32 in 1983, and 27 in 1984). The probability values are based on two-tailed *t*-tests except those for the intercepts (one-tailed).

[a] $CAR_{it}=\alpha+\beta_1 PROD_{it}+\beta_2 DISCOV_{it}+\beta_3 PURCH_{it}+\beta_4 QTYREV_{it}+\beta_5 PRICE_{it}+\beta_6 TAXCHG_{it}+\beta_7 COSTS_{it}+\beta_8 dNI_{it}+e_{it}$.

[b] $CAR_{it}=\alpha+\beta_1 PROD_{it}+\beta_2 DISCOV_{it}+\beta_3 DISCOVPP_{it}+\beta_4 PURCH_{it}+\beta_5 PURCHPP_{it}+\beta_6 QTYREV_{it}+\beta_7 PRICE_{it}+\beta_8 TAXCHG_{it}+\beta_9 COSTS_{it}+\beta_{10} dNI_{it}+e_{it}$.

[c] $CAR_{it}=\alpha+\beta_1 PROD_{it}+\beta_2 DISCOV_{it}+\beta_3 DISCOVPR_{it}+\beta_4 PURCH_{it}+\beta_5 PURCHPR_{it}+\beta_6 QTYREV_{it}+\beta_7 PRICE_{it}+\beta_8 TAXCHG_{it}+\beta_9 COSTS_{it}+\beta_{10} dNI_{it}+e_{it}$.

profits (*PROD*) increase the SM because they are no longer future outlays, and so they are a positive part of the total *TAXCHG* value. A negative coefficient for this positive part of *TAXCHG* corresponds to a tax adjustment to gross profits (which are positively related to returns). *PROD* is the largest component (in absolute mean value), so the tax adjustment to *PROD* should be a major part of *TAXCHG* and, in fact, *TAXCHG* was positive on average.[15]

Finally, in considering the negative coefficient for revisions, note that *QTYREV* is significantly negatively correlated with *PRICE*.[16] This suggests that the larger the price decline, the more positive the quantity revisions. *QTYREV* includes downward revisions that can occur when O & G prices decline and reserves are no longer economically recoverable. Thus, as prices declined during the period, firms should have had negative revisions for any reserves that dropped out of the proved category. However, quantity revisions were *larger* the more prices declined, which suggests that firms may have been motivated to adjust or manage their revisions. If firms that were doing worse had smaller returns, and if these firms also revised their estimates upward more (or downward less), this could show up as a negative coefficient for *QTYREV*. Although this is conjectural, the observed negative relation between *PRICE* and *QTYREV* is a provocative finding.

The tests of the *CSM* components were replicated with two alternative measures of *CAR*. First, an industry index was added to the market model to remove the effects of industry wide events from the return measure. The model was estimated for each firm using weekly data, and the industry index was the equally weighted return for all sample firms. *CAR*s estimated from this model should be primarily related to firm-specific events. When the dual-index model was used, the *F*-statistic for the incremental information content of model 2 compared with 1 was 2.15 (0.03 significance level). Again, this suggests that disaggregating the *CSM* into its components provides additional information. As shown in table 4, the results of the dual-index version of models 3 through 5 are similar to those for the single-index *CAR*s (table 3). One major exception is that the coefficient for *dNI* is significantly positive in the models, which suggests that net income provides information about firm-specific events affecting producers. This is in contrast to the findings with single-index *CAR*s (here and in Doran et al. 1988) that earnings did not provide information. Several *CSM* components continue to contribute to the models even after removing industry effects. This is a strong test of the information content of the components, given that they are industry-specific measures by nature and are designed to reflect trends in the O & G production economy.

Second, *CAR*s were estimated with the (single-index) market model but were measured from the beginning of April of the test year through the end of March of the following year to include the effects of the disclosure of the *CSM* data in the return measure.[17] The results of models 3 through 5 when using the April through March

---

[15] Similarly, taxes relating to *DISCOV* (*PURCH*) would be a negative part of *TAXCHG* since they decrease the SM. A negative coefficient for that part of *TAXCHG* translates into a positive effect on returns, which is the opposite effect on returns as *DISCOV* (*PURCH*).

[16] Correlation coefficients for all variables in the study are available from the author.

[17] The April through March returns are measured over a period that lags the fiscal year covered by the accounting disclosures (i.e., January through December) by three months. This lack of synchronization could obscure the relation between the accounting variables and returns.

Table 4
OLS Regression Results for Models 3 through 5: Dual-Index Returns Model

	α	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
Model 3:[a]											
Coefficient	−0.23	−1.40	−0.76	−0.81	−0.91	−0.82	−1.53	−0.25	0.59		
White's Probability	0.08	<0.001	0.02	0.07	0.005	0.01	0.002	0.41	0.01		
R-square	0.26										
Adjusted R-square	0.18										
F-statistic	3.26										
Probability	0.003										
Model 4:[b]											
Coefficient	−0.23	−1.39	−0.81	−0.45	0.13	10.66	−0.89	−0.77	−1.51	−0.24	0.58
White's Probability	0.10	0.001	0.45	0.97	0.94	0.50	0.01	0.03	0.003	0.45	0.02
R-square	0.27										
Adjusted R-square	0.16										
F-statistic	2.56										
Probability	0.011										
Model 5:[c]											
Coefficient	−0.21	−1.32	−0.69	0.73	−1.20	−0.49	−0.90	−0.93	−1.61	−0.22	0.50
White's Probability	0.13	0.001	0.10	0.60	0.04	0.35	0.005	0.07	0.002	0.48	0.06
R-square	0.27										
Adjusted R-square	0.17										
F-statistic	2.64										
Probability	0.009										

Note: Abnormal returns were estimated with a dual-index form of the market model that includes an industry index in addition to a market index. Dependent and independent variables are defined in table 1 and the text. The total number of observations is 82 (23 in 1982, 32 in 1983, and 27 in 1984). The probability values are based on two-tailed t-tests except those for the intercepts (one-tailed).

[a] $CAR_{it}=\alpha+\beta_1 PROD_{it}+\beta_2 DISCOV_{it}+\beta_3 PURCH_{it}+\beta_4 QTYREV_{it}+\beta_5 PRICE_{it}+\beta_6 TAXCHG_{it}+\beta_7 COSTS_{it}+\beta_8 dNI_{it}+e_{it}$.

[b] $CAR_{it}=\alpha+\beta_1 PROD_{it}+\beta_2 DISCOV_{it}+\beta_3 DISCOVPP_{it}+\beta_4 PURCH_{it}+\beta_5 PURCHPP_{it}+\beta_6 QTYREV_{it}+\beta_7 PRICE_{it}+\beta_8 TAXCHG_{it}+\beta_9 COSTS_{it}+\beta_{10} dNI_{it}+e_{it}$.

[c] $CAR_{it}=\alpha+\beta_1 PROD_{it}+\beta_2 DISCOV_{it}+\beta_3 DISCOVPR_{it}+\beta_4 PURCH_{it}+\beta_5 PURCHPR_{it}+\beta_6 QTYREV_{it}+\beta_7 PRICE_{it}+\beta_8 TAXCHG_{it}+\beta_9 COSTS_{it}+\beta_{10} dNI_{it}+e_{it}$.

CARs, shown in table 5, indicate a positive significant coefficient for *dNI* in models 3 and 4. This is consistent with the general finding in the accounting literature that the market appears to react to annual earnings releases, including those of O & G producers.

III. Discussion and Summary of Findings

Taken as a whole, the various analyses of the information content of the *CSM* components suggest the following general findings. First, disaggregating the *CSM* into its components, as required by SFAS No. 69, provides incremental information over reporting only the net change in reserve value. The required decomposition of the CSM is intended to provide additional information to statement users and does so. The CSM is the sum of the effects of several economic factors related to reserve value changes. To some extent, each *CSM* component is intended to reflect a different economic factor(s). Moreover, these factors can counterbalance each other. For example, production of reserves (*PROD*) could offset additions of reserves from discoveries (*DISCOV*), which would result in a combined zero net change in the SM. Thus, information could be lost if only the sum of the various effects were reported. In fact, the results indicate that knowing how and why the reserve value changed (i.e., the component composition of the *CSM*) provides additional information over just the net change. In addition, the study extends previous research by examining the entire set of *CSM* components. This reflects the nature and purpose of the *CSM* disclosure because it considers the exact set of information that is reported and also allows for more complex multivariate relationships. The findings provide a different picture of the *CSM* disclosure from that of previous studies that examined one component (e.g., discoveries) or an aggregate of components (e.g., price and quantity revisions combined).

Second, the findings for the information content of individual components were mixed. The results provide the first evidence that gross profits from production are positively associated with returns. Five other components were significantly negatively related to returns: discoveries, purchases, income taxes, quantity revisions, and price changes. Negative signs for both discoveries and purchases support arguments and evidence presented by McConnell and Muscarella (1985), Picchi (1985), and Jensen (1986a, 1986b) that the market reacted negatively to oil firms' overinvesting in E & D during the period examined, in part because of a declining trend in oil prices.

The results must be viewed with caution because of the presence of multicollinearity. Collinearity may provide additional insight into counterintuitive results for *PRICE* and *QTYREV*. Note that *PRICE* is significantly correlated with almost every variable in models 3 through 5, and *QTYREV* is correlated with *PRICE*. Further, collinearity may explain the insignificance of *COSTS*, given that they are correlated with most of the other measures. As one might expect, some of the CSM components are naturally collinear since they are based on a common set of information such as the firm's reserve holdings and price changes. Previous studies by Magliolo (1986); Harris and Ohlson (1987); Doran et al. (1988); Ghicas and Pastena (1989); and Lambert et al. (1989) found that two or more reserve value-based variables were significantly collinear. The findings in this study document the inherently collinear nature of the data for the entire CSM disclosure. Survey studies by Deakin and Deitrick (1982) and Young (1986) found that financial analysts consider the information in the CSM disclosure to be useful. However, the significant collinearity between components found in this and previous

Table 5
OLS Regression Results for Models 3 through 5: April through March Returns

	α	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}
Model 3:[a]											
Coefficient	−0.27	−1.74	−0.76	−0.55	−0.30	−0.46	−0.13	−0.53	0.67		
White's Probability	0.10	0.005	0.14	0.44	0.58	0.36	0.86	0.28	0.06		
R-square	0.25										
Adjusted R-square	0.17										
F-statistic	3.11										
Probability	0.004										
Model 4:[b]											
Coefficient	−0.20	−0.93	+0.66	23.80	−0.69	4.92	−0.01	−0.15	−0.22	−0.09	0.89
White's Probability	0.09	0.02	0.10	<0.001	0.23	0.43	0.98	0.71	0.71	0.81	0.001
R-square	0.59										
Adjusted R-square	0.53										
F-statistic	10.36										
Probability	<0.001										
Model 5:[c]											
Coefficient	−0.32	−1.35	0.01	4.31	−1.09	−1.51	−0.003	−1.36	−0.14	−0.40	0.21
White's Probability	0.05	0.02	0.99	0.01	0.16	0.02	>0.99	0.02	0.85	0.39	0.44
R-square	0.29										
Adjusted R-square	0.19										
F-statistic	2.96										
Probability	0.004										

Note: Abnormal returns are measured from April of the test year through March of the following year. Dependent and independent variables are defined in table 1 and the text. The total number of observations is 83 (24 in 1982, 32 in 1983, and 27 in 1984). The probability values are based on two-tailed *t*-tests except those for the intercepts (one-tailed).

[a] $CAR_{it}=\alpha+\beta_1 PROD_{it}+\beta_2 DISCOV_{it}+\beta_3 PURCH_{it}+\beta_4 QTYREV_{it}+\beta_5 PRICE_{it}+\beta_6 TAXCHG_{it}+\beta_7 COSTS_{it}+\beta_8 dNI_{it}+e_{it}$.

[b] $CAR_{it}=\alpha+\beta_1 PROD_{it}+\beta_2 DISCOV_{it}+\beta_3 DISCOVPP_{it}+\beta_4 PURCH_{it}+\beta_5 PURCHPP_{it}+\beta_6 QTYREV_{it}+\beta_7 PRICE_{it}+\beta_8 TAXCHG_{it}+\beta_9 COSTS_{it}+\beta_{10} dNI_{it}+e_{it}$.

[c] $CAR_{it}=\alpha+\beta_1 PROD_{it}+\beta_2 DISCOV_{it}+\beta_3 DISCOVPR_{it}+\beta_4 PURCH_{it}+\beta_5 PURCHPR_{it}+\beta_6 QTYREV_{it}+\beta_7 PRICE_{it}+\beta_8 TAXCHG_{it}+\beta_9 COSTS_{it}+\beta_{10} dNI_{it}+e_{it}$.

studies raises the question as to how financial statement users can effectively use the information, given the inherent statistical limitations of the data. This point may be useful for policymakers to consider when designing any future disclosure requirements relating to O & G producing activities.

As in any capital markets study, there could also be correlated omitted variables. Magliolo's (1986, 96) results suggested that there were additional relevant explanatory variables omitted from the model he tested. Another potential misspecification of the models relates to the expectations models used.[18] Doran et al. (1988, 394) use a random walk expectations model for net income, but do not use an expectations model for the reserve value-based measures they examine. Presumably, they assume that market participants cannot predict reserve value changes attributable to future discoveries, price changes, and revisions in the year ahead. However, Magliolo (1986) found some evidence that *DISCOV* was at least partially anticipated in the year before it was reported, but that revisions and price changes were not anticipated. On the basis of his findings, the change in *DISCOV* from the previous year was substituted for *DISCOV* in models 3 through 5 but the nature of the results did not change.

Third, the tests with dual-index *CAR*s indicated that the *CSM* components for production, discoveries, purchases, quantity revisions, price change, and income taxes provide information even after controlling for industry effects. Finally, net income was positively related to returns both when a dual-index returns model was used (January through December) and when returns were measured from April through March (single-index model). This is in contrast to the findings in this study and Doran et al. (1988) that net income did not provide information when returns were estimated with a single-index, January through December model.

---

[18] Lanen and Thompson (1988) find that the sign of association between abnormal returns and firm characteristics depends on details of the market setting, such as the information set of various market participants. They show that the effects of market characteristics can reverse the predicted sign of association.

Appendix A

Replication of Analyses of Doran et al. (1988)

Doran et al. (1988, 392) conducted separate tests for refiners in addition to producers, but point out that, given that exploration and production is a smaller portion of the total activities of refiners, the reserve measures for refiners are likely to exhibit lower associations with returns than those for producers. Consequently, the replication includes only producers to facilitate the detection of the relation between the SFAS No. 69 reserve value-based disclosures and returns. See Doran et al. (1988, 392–99) for a discussion of their research method.

Model 2 through 2c in Doran et al. (1988, 392–93) were replicated using SFAS No. 69 data over the same period, 1982–1984, but with a different sample of producers (see section II). The variables in the replication were measured according to the definitions in Doran et al. except that abnormal returns were calculated by accumulating weekly (vs. monthly) prediction errors from the market model over the fiscal year.

A major difference between the replication and Doran et al. (1988) is the exclusion of Canadian and other foreign firms, as well as limited partnerships and royalty trusts, from the replication sample to control for economic differences between these entities and U.S. corporations that could affect returns (e.g., tax laws). Previous oil and gas research provides evidence that including Canadian firms can potentially confound results relating to U.S. firms (see Collins and Dent 1979). Sixteen of the 88 observations in the Doran et al. sample (about 18 percent) are Canadian firms. Three additional Doran et al. observations are excluded because the firms provided RRA data rather than SFAS No. 69 data in 1982, and another three observations are excluded because the relevant firm is not primarily a producer. There are 25 other observations in the Doran et al. sample that are not in the replication, and 38 observations in the replication that are not in Doran et al. When the replication was repeated with a subsample that included the common observations in the two samples, the results were unchanged.

The replication results shown below are very similar to those of Doran et al. (1988, 409), although revisions had a positive, significant coefficient in Doran et al., whereas it is negative and insignificant in the replication. The replication regressions also exhibit higher explanatory power, which suggests that excluding Canadian firms eliminated some of the noise in the estimated regressions. The replication helps reconcile Doran et al.'s results, which examine a few aggregated measures, with those found when the entire set of (disaggregated) SFAS No. 69 CSM components is considered.

Replication Results, Pooled Cross-Sectional OLS Regressions

Model 2:	α_0	α_3	α_4	β_1	β_2	*F-statistic (Probability)*	*Adjusted R-square*
Coefficient	−0.36	0.95	−0.04	0.55		23.79	0.46
White's Probability	0.01	<0.001	0.79	0.13		(<0.001)	
Model 2a:							
Coefficient	−0.37	0.94	−0.05	0.55	0.10	17.66	0.45
White's Probability	0.01	<0.001	0.74	0.14	0.64	(<0.001)	
Model 2b:							
Coefficient	−0.36	0.93	−0.04	0.54	−0.07	17.74	0.45
White's Probability	0.01	<0.001	0.79	0.14	0.42	(<0.001)	
Model 2c:							
Coefficient	−0.36	0.95	−0.04	0.54	0.02	17.62	0.45
White's Probability	0.01	<0.001	0.79	0.14	0.92	(<0.001)	

For each of the models, the intercept and coefficient names are those used in Doran et al. (1988, 392–93). Probability levels given for the α and β coefficients are based on one and two tailed t-tests, respectively, and the probability values are based on White's (1980) covariance estimator. The total number of observations is 82 (23 in 1982, 32 in 1983, and 27 in 1984).

Appendix B
Calculation of the CSM Components

SFAS No. 69 (par. 33) specifies that the price change component be computed by applying current year-end prices to the *previous* year's reserve estimates after subtracting current-year production. That is, this component should reflect only the net change in prices and costs relating to wells that were proved at the beginning of the year and that were not produced. The quantity revisions component is calculated by applying current year-end sales prices and costs to the quantity revisions for the year. Also, for reserves to be classified as "proved," they must be recoverable under existing economic conditions. Consequently, quantity revisions may include revisions resulting from changes in the economic recoverability of reserves, such as downward revisions that can occur when sales prices for oil and gas decline. The component for the net change in future income taxes represents the change in discounted future income taxes (relating to future cash flows) from the previous balance sheet date to the current balance sheet date. This change implicitly includes the income taxes incurred during the current period. Thus, the other components are on a pretax basis, including the "accretion of discount" component, which is calculated by multiplying beginning discounted future cash flows before income taxes by 10 percent. By definition, the components must sum up exactly to the CSM. Any discrepancy between the CSM and the sum of the first ten components is netted out by the last component, "other" (i.e., it is a plug figure).

To understand the nature of the components, one must recognize that each measures the CSM attributable to that component. For example, if (discounted) income taxes relating to proved reserves *increase*, the SM *decreases* so the income tax component would be negative. Furthermore, because the components are derived from the current year's and previous year's SMs, they reflect the assumptions used in the SM calculation. For example, the component for purchases (sales) of minerals in place reflects the CSM attributable to purchases (sales). Its value does not necessarily equal the price paid (received) for the minerals because the assumptions required for the SM (e.g., year-end prices and costs) may not coincide with those implicit in the purchase (sales) price.

References

Alciatore, M. L. 1990. The reliability and relevance of reserve value accounting data: A review of the empirical research. *Journal of Accounting Literature* 9: 1–38.

Arthur Andersen & Co. 1983. *Oil and Gas Reserve Disclosures: 1980–1983*. Houston, TX: Arthur Andersen & Co.

Barrett, M. E. 1989. The demise of Gulf Oil Corporation. Case study, Southern Methodist University, Dallas, TX.

Beaver, W., P. Griffin, and W. Landsman. 1982. The incremental information content of replacement cost earnings. *Journal of Accounting & Economics* 4 (July): 15–39.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley.

Bernard, V. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* 25 (Spring): 1–48.

Christie, A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting & Economics* 9: 231–58.

Collins D., and W. Dent. 1979. The proposed elimination of full cost accounting in the extractive petroleum industry: An empirical assessment of the market consequences. *Journal of Accounting & Economics* 1: 3–44.

Deakin, E. B. 1979. An analysis of differences between non-major oil firms using successful efforts and full cost methods. *The Accounting Review* 54 (October): 722–34.

———, and J. W. Deitrick. 1982. An evaluation of RRA and other supplemental oil and gas disclosures by financial analysts. *Journal of Extractive Industries Accounting* 1 (Spring): 63–70.

Doran, B. M., D. Collins, and D. Dhaliwal. 1988. The information of historical cost earnings relative to supplemental reserve-based accounting data in the extractive petroleum industry. *The Accounting Review* 63 (July): 389–413.

Financial Accounting Standards Board. 1982. *Statement of Financial Accounting Standards No. 69: Disclosures About Oil and Gas Producing Activities*. Stamford, CT: FASB.

Ghicas, D., and V. Pastena. 1989. The acquisition value of oil and gas firms: The role of historical costs, reserve recognition accounting, and analysts' appraisals. *Contemporary Accounting Research* 6 (Fall): 125–42.

Harris, T., and J. Ohlson. 1987. Accounting disclosures and the market's valuation of oil and gas properties. *The Accounting Review* 62 (October): 651–70.

Harvard Business School. 1984. Gulf Oil Corp.—Takeover. Case study. Harvard Business School, Boston, MA.

Imhoff, E., R. Lipe, and D. Wright. 1992. The effects of recognition versus disclosure on shareholder risk and executive compensation. Working paper, University of Michigan, Ann Arbor, MI.

Jensen, M. 1986a. Agency costs of free cash flow, corporate finance, and takeovers. *American Economics Association Papers and Proceedings* (May): 323–29.

———. 1986b. The takeover controversy: Analysis and evidence. Working paper No. 86–01, University of Rochester, Rochester, NY.

Lambert, R., D. Larcker, and J. Magliolo. 1989. An econometric analysis of incremental information content studies. Working paper, University of Pennsylvania, Philadelphia, PA.

Lanen, W., and R. Thompson. 1988. Stock price reactions as surrogates for the net cash flow effects of corporate policy decisions. *Journal of Accounting & Economics* 10 (December): 311–34.

Lipe, R. 1986. The information contained in the components of earnings. *Journal of Accounting Research* 24 (Supplement): 37–64.

McConnell, J., and C. Muscarella. 1985. Corporate capital expenditure decisions and the market value of the firm. *Journal of Financial Economics* 14 (September): 399–422.

Magliolo, J. 1986. Capital market analysis of reserve recognition accounting. *Journal of Accounting Research* 24 (Supplement): 69–108.

Picchi, B. 1985. *Structure of the U.S. Oil Industry: Past and Future*. Salomon Brothers (July).

Securities and Exchange Commission. 1979. *Accounting Series Release No. 269: Oil and Gas Producers—Supplemental Disclosures on the Basis of Reserve Recognition Accounting*. Washington, D.C.: Securities and Exchange Commission.

———. 1982. *Financial Reporting Release No. 9: Supplemental Disclosures of Oil and Gas Producing Activities*. Washington, D.C.: Securities and Exchange Commission.

Spear, N. 1992. The incremental information in SFAS No. 69 reserve disclosures over historical cost income and its cash and accrual components. Working paper, Flinders University of South Australia.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 40 (May): 817–38.

Young, G. 1986. An investigation of the usefulness of the financial statements as supplemented by oil and gas producing activities disclosures. Ph.D. dissertation, George Washington University, Washington, D.C.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 657–667

Organizational Form Choice and the Valuation of Oil and Gas Producers

Wayne H. Shaw
University of Colorado at Boulder
Heather A. Wier
University of Alberta

SYNOPSIS AND INTRODUCTION: The Tax Reform Act of 1986 reduced individual income tax rates below that of corporations. The fear that this would lead to a systematic disincorporation has not apparently materialized since the major stock exchanges report that only about 100 partnerships were traded during the next five years.[1] Scholes and Wolfson (1992) suggest that this result is predictable because of the additional nontax costs of operating as a partnership, which include increased transactions costs, more restricted access to capital markets, and less control over management. Guenther (1992) and Terando and Omer (1992) provide evidence that firms must have considered both tax and nontax costs when choosing organizational form.

This study examines whether the factors taken into consideration in organizational form choice also affected the market value of a sample of firms in the oil and gas industry during the period 1985–1988. We chose for our tests a valuation model used by Harris and Ohlson (1987) and others[2] since many of the publicly traded (master limited) partnerships (MLPs) were created in the oil industry.

The studies by Guenther (1992) and Terando and Omer (1992) show that MLPs generally had only one line of business, less debt, and higher dividend yields than their corporate counterparts. These differences are predictable in view of the tax consequences of operating in the partnership form. Guenther also finds MLPs to be less profitable, which is consistent with higher nontax costs of the partnership form. We show that MLPs in-

---

[1] An example of the concerns can be found in Murray and Langley (1987).

[2] Other studies include Magliolo (1986), Doran et al. (1988), and Ghicas and Pastena (1989). A summary of the literature can be found in Alciatore (1990).

We appreciate comments from three anonymous referees and participants in workshops at Cornell University, the University of Colorado at Boulder, the University of Maryland, and the University of Connecticut. We also appreciate financial support from the University of Colorado at Boulder and the Johnson Graduate School of Management Accounting Institute.

Submitted January 1991.
Accepted February 1992.

vested significantly less in exploration for new deposits. The combination of the lower exploration expenditures, higher dividend yields, and poor financial performance suggests that MLPs may have been set up as limited-life entities to distribute assets to their unit holders in a tax-efficient manner.[3]

We extend the Harris and Ohlson (1987) valuation model by adding dividends and exploration levels. Exploration levels are found significant for both MLPs and corporations, but dividends are relevant only in the MLP model. We also find that dividend levels are significantly explained by asset values only for the MLPs. These valuation differences are consistent with tax-motivated organizational form choice and the perceived passive nature of the MLPs.

Key Words: *Organizational form, Dividends.*

Data Availability: *The data are available from the authors upon request.*

THIS article is structured as follows. In section I, we describe the sample selection process and some comparative statistics on the two groups of firms. A discussion of potential differences in market perception based on organizational choice and calculation of test variables are provided in section II. Results of regression models of imputed market values and dividend levels are shown in section III and IV. Sensitivity tests of results are given in section V. Conclusions are presented in the last section.

I. Sample Description

We begin our examination in 1985 because most MLPs were created in the 1985–1986 period. The sample is limited to firms from the 1311 SIC category, Crude Oil and Natural Gas, that had publicly available equity market prices during 1985–1988. The 1311 SIC category is convenient because the Financial Accounting Standards Board's *Statement of Financial Accounting Standards No. 69: Disclosures About Oil and Gas Producing Activities* (FASB 1982) required oil and gas firms to provide supplementary oil and gas valuation disclosures that are relatively consistent across firms. During the 1985–1988 period, we found 26 MLPs and 41 corporations meeting our restrictions.[4]

The primary difference between MLPs and corporations is that MLPs are not subject to an entity level tax. Guenther (1992) and Terando and Omer (1992) focus on this difference in examining the debt and dividend policies of the two groups of firms. Both studies report results consistent with the predictions that MLPs have lower debt levels because of the loss of the corporate interest tax deduction and higher dividend payout levels because of the absence of double taxation on dividends.

Table 1 provides some descriptive data on the MLP and corporate oil and gas producers for 1986, the first full year of operation for many of our MLPs.[5] Like Guenther

---

[3] David M. Leuschen, a partner with Goldman, Sachs & Co., describes an MLP as "somewhat of a bond equivalent." Tell (1985, 1986) asserts that an MLP is valued according to its dividend yield.

[4] The small sample size for both the MLPs and corporations increases concern over the effects of extreme observations. The results of the diagnostic tests (presented later) indicate that collinearity and heteroskedasticity are not problems in this study.

[5] Similar results are found for the other three years.

Table 1
Comparison of Selected Accounting Numbers and Ratios of MLPs and Corporate Oil and Gas Producers in the 1311 SIC Code

(values from 1986 annual reports)

Accounting Number or Ratio	*MLPs* (n=23)	*Corporations* (n=41)	*t-statistic* (*Significance Level*)
Total Assets (in millions)	572.6481	850.4683	−0.5871 (0.5595)
Debt/Total Assets	0.3565	0.4572	−1.6384 (0.1067)
Long-Term Debt/Common Equity	0.3233	0.7452	−2.1891 (0.0325)
Non-Oil and Gas Assets/Total Assets	0.1239	0.3442	−6.9113 (0.000)
ROA	−0.2855	−0.0483	−3.6023 (0.0014)
YIELD	0.1030	0.0061	4.8485 (0.0001)
PPAYOUT	75.6%	4.3%	—
ACQ	0.5260	0.5808	−0.2248 (0.8229)
EXP	0.0969	0.2059	−2.5621 (0.0129)
DEV	0.2070	0.2987	−1.1349 (0.2612)
BETA	0.3464	0.6670	−1.8233 (0.0744)

Note: ROA = net income/total assets,
YIELD = yearly dividend rate/price per unit,
PPAYOUT = percentage of firms whose dividends exceeded net income,
ACQ = yearly acquisition costs/equivalent barrels of proved reserves,
EXP = yearly exploration costs/equivalent barrels of proved reserves,
DEV = yearly development costs/equivalent barrels of proved reserves, and
BETA = market beta provided by CRSP at the end of each year under examination.

Mean values reported.

(1992), we find no significant size difference, but find that MLPs had on average a lower debt-to-asset ratio (significant at the 0.11 level, two-tailed).

The debt-to-asset ratio is, however, somewhat misleading because it aggregates permanent long-term debt with short-term payables. We, therefore, examine ratios of long-term debt to common equity, which averaged 32.33 percent for MLPs and 74.52 percent for corporations. This difference is significant and consistent with both the absence of an interest tax deduction and the likely restrictive access to capital markets for MLPs.

Another tax-induced characteristic of MLPs is that they are formed with a single line of business.[6] Non-oil and gas assets for MLPs averaged only 12.4 percent of total assets, compared to 34.4 percent for corporations, which also had other lines of business more often.

---

[6] The motivation for a single line of business constraint was strengthened by the Revenue Act of 1987, which revoked partnership status for any existing MLP that added a new line of business after December 17, 1987. The Act also denies partnership status to any new MLP unless the entity earns 90 percent or more of its income from interest, dividends, rents, or natural resources production.

Like Guenther (1992), we find differences in firm profitability conditional on organizational form choice. Although both groups were negatively affected by low oil prices during the test period, with prices averaging about \$13 a barrel compared to \$28 in the 1981–1984 period, the MLPs had significantly weaker performance; an average 28.55 percent rate of return on assets compared to −4.83 percent for corporations. Although much of the difference in 1986 was attributable to asset write-downs by the MLPs, they still showed an average −1 percent return on assets for the other three years, compared to 1 percent for the corporations.

Finally, MLPs and corporations differed on dividend yields, asset reinvestment levels, and market beta. The average MLP dividend yield of 10.3 percent was significantly higher than the 0.6 percent yield for corporations. MLP dividends were also high relative to earnings levels; 75 percent of the MLPs paid dividends in excess of earnings, while only 4.3 percent of the corporations did.

We examine three measures of asset reinvestment contained in the SFAS No. 69 disclosures: yearly property acquisition costs, development costs, and exploration costs. Yearly oil and gas property acquisition cost and development costs were not significantly different, but MLPs purchased primarily proved reserves rather than undeveloped properties, as evidenced by the significantly higher oil and gas exploration costs of the corporations and by the yearly increases in proved and unproved reserves.

MLPs also had significantly lower market betas than corporations, which, combined with high payout levels and lower exploration costs, is consistent with arguments in the popular press that MLPs are passive investments whose return processes are less affected by general market movements.[7]

II. Development of Tests

The basic oil and gas valuation model used by Magliolo (1986), Harris and Ohlson (1987, 1990), and others, documents links between the market value of oil and gas assets and various financial statement valuations of the assets, including book value and three cash flow measures disclosed under SFAS No. 69. However, none of these financial disclosures explicity captures firm growth opportunities and dividend payout levels, both viewed as value-relevant in the finance literature (see, e.g., Shapiro 1990). Since MLPs and corporations differ on both of these aspects, we add to the basic model a dividend levels variable and exploration costs as a surrogate for growth potential.[8]

The expanded valuation model we use is as follows:

$$IV = a + b_1 BVSE + b_2 BVFC + b_3 PV + b_4 EXP + b_5 DIV + e, \qquad (1)$$

where,

IV=(market value of common equity+book value of liabilities−book value of non-oil and gas properties)/equivalent barrels of proved reserves,

---

[7] Hamada (1972) provides an alternative explanation that links lower betas with lower debt levels.

[8] We also added property acquisition and asset development costs to the model. Neither has added significant explanatory power.

BV = book value of oil and gas properties/equivalent barrels of proved reserves,
PV = before-tax present value of net future cash flows from proved reserves/equivalent barrels of proved reserves,
EXP = yearly exploration costs incurred/equivalent barrels of proved reserves,
DIV = yearly dividends per share × the number of shares outstanding/equivalent barrels of proved reserves, and
e = an error term.

The market value of common equity was calculated by using closing market prices three and a half months after the end of the relevant fiscal year. As in Harris and Ohlson (1987), the three and a half month lag was intended to capture any effects on market value from the release of annual earnings and supplementary reserve value disclosures. Shares outstanding were collected from disclosures in the first quarterly report. We used book value for both liabilities and unrelated assets.

The *BV* and *PV* variables have been the primary focus of prior studies. We collected the book value of oil and gas properties from the balance sheet and footnote disclosures of the firms' annual reports. Any fixed assets not directly part of the costs of oil and gas properties, such as plant facilities and equipment, were considered non-oil and gas assets. In conformance with Harris and Ohlson (1987), *BV* includes costs being currently amortized, costs not being currently amortized, and cumulative amortization.

The sample firms reported their oil and gas operations with either successful efforts (SE) or full cost (FC) accounting methods. Under FC, all exploration costs are capitalized, but SE requires capitalization of only the costs of productive wells. Of the 41 (26) corporate (MLP) oil and gas producers, 29 (15) used the FC method in 1985. During 1986, four corporations and two MLPs switched to SE. Because of the potential differences in *BV*, all regressions were estimated with separate book value measures based on the accounting method employed.

We obtained measures for the *PV* variable from the supplementary SFAS No. 69 disclosures. *PV* is calculated as the discounted value of the yearly cash flows from proved oil and gas reserves. SFAS No. 69 mandates that a 10 percent discount rate be used in the calculation and that the present value measure be reported by corporations on an after-tax basis. MLPs report before-tax levels since they incur no entity level taxation. Prior studies, such as Harris and Ohlson (1987), examined *PV* on a before-tax basis, which was required under *Accounting Series Release No. 253* (SEC 1978). To be consistent with these studies, we converted the present-value measure on corporations to a before-tax basis by adding back the taxes reported in the footnote disclosure under SFAS No. 69. However, all inferences are substantially identical to those obtained by using the reported after-tax numbers.

The primary finding of Harris and Ohlson (1987) is that *BV* dominates *PV* in explaining corporate *IV*. This result seems inconsistent with the FASB's expressed concerns when enacting SFAS No. 69. The FASB believed the supplementary disclosures were necessary since book value could include substantial measurement error with the use of historical costs and fixed rate amortization schedules. Harris and Ohlson conjecture that their result may be attributable to the market's placing credence in historical cost accounting because of its objectivity. An alternative explanation for their result is that *BV* includes costs of proved and unproved reserves, while *PV* contains only proved reserves in its calculation. We examine whether the result of Harris and Ohlson for 1979–1983 continues to hold during 1985–1988 by testing the following hypothesis:

H1: Book values continue to dominate present values in explaining market values for both MLPs and corporations.

If historical cost is preferred by the market, the MLP valuations should be closely related to book value since 15 of the 26 MLPs wrote up their assets to "market value" at the acquisition date.[9] Since not all firms use the same accounting methods and amortization periods, the relationship between *BV* and *IV* should be strongest in the *earlier* years.

There are reasons to believe, however, that the result found for corporations may not hold for MLPs. First, since MLPs have primarily proved reserves and less exploration activities, there is less concern that the market may be placing some value on unexplored and undeveloped assets contained in *BV* but not *PV*. Second, the financial press often stressed a link between reserves and market value for MLPs when discussing reasons for their creations, viewing an MLP merely as a payout vehicle for previously discovered oil.

The final two variables, *EXP* and *DIV*, are added for the reasons cited earlier and because MLPs appear to have been viewed as high-yield bond equivalents rather than as going concerns. On the basis of this reasoning, we examine whether:

H2: *EXP* is significant in explaining the market value of both MLPs and corporations.

H3: *DIV* is more important to the valuation of MLPs than corporations.

Table 1 and the popular press suggest that MLPs would, on average, be expected to produce out of current production and purchased proved reserves to maintain their high dividend yields. An MLP that incurs significant exploration costs provides a signal inconsistent with a view of an MLP as a limited-life dividend payout vehicle. This signal is expected to have a significant effect on MLPs' valuation. In contrast, corporations are expected to invest in new growth opportunities. Therefore, a low level of reinvestment would be viewed as a signal that the firm either did not have or was unable to take advantage of growth opportunities. These arguments would suggest a positive relationship for both types of firms between market value and exploration levels.

DIV represents the expected dividend payout level. It is computed as four times the first quarter dividend. This calculation assumes that the firm has not disclosed to the market any plan to change dividend levels in the future, and captures the extent to which oil and gas producers are viewed as passive investments with market value based on projected dividend streams. This variable is potentially more important in the valuation of MLPs because of their higher dividend levels.[10] Finally, in conformance with prior studies (see, e.g., Harris and Ohlson 1987), all variables are scaled by the equivalent barrels of proved reserves to reduce potential problems with heteroskedasticity.[11]

[9] The firms did not disclose how adjustments to market were conducted. We contacted a Big Six partner who was involved in four of the transactions. He stated that the amount allocated to oil and gas properties was equal to the purchase price less market values allocated to non-oil and gas assets to avoid booking goodwill for MLPs.

[10] Shaw (1991) purports that MLP dividend levels are strongly linked to market returns. He finds that the announcement of a cut in dividends leads to a decline in unit price by approximately 8 percent. This reaction is about twice the amount found for corporations by Aharony and Swary (1980). Our emphasis is on dividend levels because of discussions in the popular press.

[11] Throughout, we used the standard gas/oil conversion that equates 6,000 cubic feet of gas to one barrel of oil. Prior results (see, e.g., Harris and Ohlson 1987) have been shown to be insensitive to the conversion rate used.

III. Explanation of Imputed Market Value

As shown in table 2, the regression results differ substantially by the organizational form choice.[12] The book value variables are important in explaining *IV* for the MLPs only in 1985. The dividend level is found to be important in explaining the market value of MLPs. Also, the supplementary disclosures of the *PV* and *EXP* seem to be important explanatory variables, with significant coefficients in 1986 for *EXP* and in 1987 for both.

In contrast, the results for corporations are consistent with those of Harris and Ohlson (1987) in that (1) the book value measures are significant in explaining the market value of oil and gas reserves and (2) the coefficient on the SE accounting method is higher than that for the FC method. Only *EXP* adds (in three or four years) significant incremental explanatory power to the model. The coefficients on the *PV* and the *DIV* variables are not significant in any of the four years.

These results suggest several observations. First, the learning hypothesis proposed by Beaver et al. (1982) is not supported. Second, the valuation models of MLPs and corporations are different in that the dividend variable plays an important role only in the MLP valuation, which is consistent with the tax motivation for MLPs. Third, unlike corporations, *PV* is not dominated in the MLP valuation. Finally, exploration costs appear to capture a growth element not contained in either *BV* or *PV* in firm valuation for both groups.

Two *caveats* are pertinent. The 1985 results might have been affected by the dramatic oil price movements in early 1986, and the 1988 results may be affected by the weakening financial viability of many MLPs. By the end of 1989, 12 of the 26 MLPs had been liquidated. Of the remaining 14 firms, four omitted dividends in early 1989.

IV. Explanation of Dividend Levels

The previous results suggest examining how the firms set dividend levels. We examine this issue by estimating the following model:

$$DIV = a + b_1 BVSE + b_2 BVFC + b_3 PV + b_4 EXP. \tag{2}$$

Prospectuses issued by many of the MLPs described dividend levels as a function of current and future purchases of production capability and development of properties. Current production capacity could be estimated either by the book value or the supplementary present value of the oil and gas properties. Also, as discussed in the last section, it appears that *EXP* captures a growth element in market value. If dividends are set on the basis of expected future growth, *EXP* could also explain this cross-sectional variability. The relatively lower level of dividend payments for corporations suggests that the links between dividends and asset values would be weaker than for MLPs.[13] We thus test the following hypothesis:

H4: Dividend levels are established on the basis of the value of the oil and gas reserves and growth opportunities as evidenced by exploration costs.

---

[12] We do not present tests of two supplementary disclosures, net future cash flows and direct profit margin, which were included in some other studies. We reestimated all the regressions we report including each and both of these variables. As in Harris and Ohlson (1987), these variables did not add any explanatory power to the equations reported. Since all independent variables are expected to be positively related to market value, the *t*-statistics reported are one-tailed.

[13] Smith (1985) demonstrates the effect that a change to MLP organizational form often has on firm dividend policy. Mesa Petroleum, for example, increased its quarterly dividend level from five cents to fifty cents a share immediately upon becoming a partnership.

Table 2
Multiple Regression Results of Imputed Value for the Years 1985–1988
(*t*-statistics in parentheses)

	Intercept	*BV—SE*	*BV—FC*	*PV*	*DIV*	*EXP*	$\bar{R}^2$	N
MLPs								
1985	4.087	0.489	0.487	−0.385	3.206	2.261	0.714	16
	(2.057)*	(1.496)	(2.379)*	(1.156)	(3.311)*	(0.631)		
1986	0.864	0.258	0.044	0.606	1.430	7.798	0.732	23
	(0.683)	(0.876)	(0.144)	(1.444)	(1.535)	(1.904)*		
1987	1.724	−0.458	−0.448	0.776	2.328	11.994	0.890	22
	(1.775)	(1.733)	(1.969)	(2.379)*	(4.339)*	(3.597)*		
1988	2.130	−0.110	0.668	0.297	2.730	1.535	0.752	20
	(1.372)	(0.335)	(0.238)	(0.772)	(4.077)*	(0.332)		
Corporations								
1985	1.134	0.903	0.399	0.073	−9.432	1.679	0.421	38
	(0.664)	(2.999)*	(1.917)*	(0.240)	(1.526)	(2.199)*		
1986	6.805	1.202	0.226	0.371	−7.156	5.066	0.511	37
	(0.493)	(3.945)*	(0.929)	(0.880)	(1.167)	(2.134)*		
1987	2.575	0.824	0.506	−0.108	−4.696	6.376	0.225	41
	(1.223)	(2.533)*	(1.782)*	(0.231)	(0.712)	(2.355)*		
1988	2.777	1.030	0.827	−0.147	−4.299	−2.379	0.346	40
	(2.183)*	(4.040)*	(3.849)*	(0.496)	(1.325)	(1.558)		

Note: *IV*=(market value of equity+book value of liabilities−book value of non-oil and gas properties)/equivalent barrels of proved reserves. Year is the dependent variable,
BV=book value of oil and gas properties / equivalent barrels of proved reserves. SE is for successful efforts firms and FC is for firms using full costing,
PV=present value of net future cash flows / equivalent barrels of proved reserves,
DIV=(yearly dividend rate × number of shares outstanding) / equivalent barrels of proved reserves, and
EXP=(yearly exploration costs) / equivalent barrels of proved reserves.

* Significant at the 0.05 level, one-tailed.

The results of estimating model 2 are presented in table 3. For the four years, about half of the variation in *DIV* for MLPs is explained by the supplementary present value of the proved oil and gas reserves and the growth element measured by exploration costs. Book value is not significant in any year.[14]

V. Potential Regression Model Issues

Several studies (e.g., Ghicas and Pastena 1989; Magliolo 1986) discuss the potential effects of multicollinearity, heteroskedasticity, and extreme observations on model estimation.[15] We constructed two tests to examine the effect of multicollinearity on the

---

[14] We also estimated the equation for the one-third of the corporations that paid dividends. The estimation of the model provides no meaningful explanation of how corporate dividend levels are set.

[15] Lys (1986) also discusses the potential problems that result from measurement error in the dependent variable.

Table 3
Multiple Regression Results of Dividend Levels for the Years 1985–1988
(*t*-statistics in parentheses)

	Intercept	*BV—SE*	*BV—FC*	*PV*	*EXP*	$\bar{R}^2$	N
MLPs							
1985	−0.726	0.052	0.036	0.183	1.989	0.438	16
	(1.256)	(0.515)	(0.574)	(2.081)*	(1.982)*		
1986	−0.491	−0.817	−0.150	0.286	3.109	0.681	23
	(1.642)	(1.134)	(2.153)	(3.488)*	(4.246)*		
1987	−0.522	−0.169	−0.217	0.358	3.848	0.565	22
	(1.243)	(1.943)	(2.078)	(3.001)*	(3.252)*		
1988	−0.807	−0.081	−0.164	0.352	2.194	0.556	20
	(1.439)	(0.647)	(1.649)	(3.006)*	(1.299)		
Corporations							
1985	−0.040	0.022	−0.000	0.016	0.038	0.023	15
	(0.498)	(0.812)	(0.018)	(0.969)	(0.028)		
1986	0.119	0.075	0.050	−0.061	0.009	−0.046	11
	0.759	(1.522)	(1.156)	(1.141)	(0.036)		
1987	0.158	0.061	0.035	−0.033	0.690	0.030	14
	(1.779)	(1.723)	(1.173)	(1.008)	(1.263)		
1988	0.099	0.081	0.045	−0.039	−0.071	0.080	16
	(0.956)	(2.120)*	(1.411)	(1.067)	(0.272)		

Note: *DIV*=(yearly dividend rate × number of shares outstanding)/equivalent barrels of proved reserves. Year is the dependent variable,
BV=book value of oil and gas properties/equivalent barrels of proved reserves. SE is for successful efforts firms and FC is for firms using full costing,
PV=present value of net future cash flow/equivalent barrels of proved reserves, and
EXP=yearly exploration cost/equivalent barrels of proved reserves.

* Significant at the 0.05 level, one-tailed.

interpretation of the *t*-statistics. First, we calculated diagnostic condition indices suggested by Belsley et al. (1980). A high condition index (exceeding 0.5) indicates that the data is ill-conditioned. The procedure then calculates the proportion of the variance of the estimate associated with each principal component. A collinearity problem occurs when a component associated with a high condition index contributes strongly to the variance of two or more variables. We find that a potential problem exists only in 1986, when the principal component with a high condition index contributed beyond the 0.5 cutoff to the variance of *PV* and the two *BV* measures for MLPs.

Second, we calculated Lagrange multiplier statistics that express the effects of each variable on market value with all other variables constrained to zero.[16] These statistics indicate that the results reported in tables 2 and 3 are not driven by collinearity between the regressors. On the basis of these tests, we conclude that multicollinearity is not an issue in this study.

[16] See Bhagat and Jefferis (1991) for a discussion of the procedure.

To test for heteroskedasticity, we calculated a chi-squared statistic suggested by White (1980). The statistic was not significant for any of the regressions reported in tables 2 and 3.

We also examined the effect of extreme observations on the results, using the approach of Belsley et al. (1980). Eliminating influential observations did not affect the interpretations of results in tables 2 and 3. Finally, the *t*-statistics were reestimated by using a boot-strapping technique described in Noreen (1989). These *t*-statistics, which are distribution-free, were found to be similar to those reported in tables 2 and 3.[17]

VI. Conclusions

In summary, we found that partnerships (MLPs) in the oil and gas industry (1) had lower debt-to-equity ratios and higher dividend yields and payout ratios than did corporations, (2) were less profitable than corportions, (3) invested less in exploration activities, and (4) had lower market betas than corporations. 'These results are consistent with assertions in the popular press that both tax and nontax costs affect firm organizational form choice.

---

[17] A recent study has questioned the appropriateness of the OLS multiple regression model. Lambert et al. (1991) used a latent variable approach in estimating the valuation equation. They demonstrate explanatory power for both book value and supplementary reserve value-based disclosures.

References

Aharony, J., and I. Swary. 1980. Quarterly dividend and earnings announcements and stockholders' returns, an empirical analysis. *Journal of Finance* 36 (March): 1–12.

Alciatore, M. L. 1990. The reliability and relevance of reserve value accounting data: A review of empirical research. *Journal of Accounting Literature* 9: 1–38.

Beaver, W. H., P. A. Griffin, and W. R. Landsman. 1982. The incremental information content of replacement cost earnings. *Journal of Accounting & Economics* 4 (July): 15–39.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York: Wiley.

Bhagat, S., and R. H. Jefferis. 1991. Voting power in the proxy process. *Journal of Financial Economics* 30 (November): 193–225.

Doran, B. M., D. W. Collins, and D. S. Dhaliwal. 1988. The incremental information content of replacement cost earnings. *The Accounting Review* 63 (July): 389–412.

Financial Accounting Standards Board. 1982. *Statement of Financial Accounting Standard No. 69: Disclosures About Oil and Gas Producing Activities*. Stamford, CT: FASB.

Ghicas, D., and V. Pastena. 1989. The acquisition value of oil and gas firms: The role of historical costs, reserve recognition accounting, and analysts' appraisals. *Contemporary Accounting Research* 6 (Fall): 125–42.

Guenther, D. A. 1992. Taxes and organizational form: A comparison of corporations and master limited partnerships. *The Accounting Review* 67 (January): 17–45.

Hamada, R. S. 1972. The effect of the firm's capital structure on the systematic risk of common stocks. *Journal of Finance* 27 (May): 435–52.

Harris, T. S., and J. A. Ohlson. 1987. Accounting disclosures and the market's valuation of oil and gas properties. *The Accounting Review* 62 (October): 651–70.

——, and ——. 1990. Accounting disclosures and the market's valuation of oil and gas properties: Evaluation of market efficiency and functional fixation. *The Accounting Review* 65 (October): 764–80.

Lambert, R. A., D. F. Larcker, and J. Magliolo. 1991. An econometric analysis of incremental information content studies. Working paper, University of Pennsylvania, Philadelphia.

Lys, T. 1986. Discussion of capital analysis of reserve recognition accounting. *Journal of Accounting Research* 24 (Supplement): 109–11.

Magliolo, J. 1986. Capital market analysis of reserve recognition accounting. *Journal of Accounting Research* 24 (Supplement): 69–108.

Murray, A., and M. Langley. 1987. Relatively new form of business structure is causing controversy. *Wall Street Journal* (June 30): 1.

Noreen, E. 1989. *Computer-Intensive Methods for Testing Hypotheses: An Introduction*. New York: Wiley.

Scholes, M. S., and M. A. Wolfson. 1992. *Taxes and Business Strategy: A Planning Approach*. Englewood Cliffs, NJ: Prentice Hall.

Securities and Exchange Commission. 1978. *Accounting Series Release No. 253: Adoption of Requirements for Financial Accounting and Reporting Practices for Oil and Gas Producing Activities*. Washington D.C.: U.S. Government Printing Office.

Shapiro, A. C. 1990. *Modern Corporate Finance*. New York: Macmillan.

Shaw, W. H. 1991. Master limited partnerships: An examination of changes in dividend policy. *Contemporary Accounting Research* 8 (Spring): 407–23.

Smith, G. 1985. The master at work. *Barron's* 63 (October 7): 16–22.

Tell, L. J. 1985. Hottest thing in oil. *Barron's* 63 (October 7): 16–22.

——. 1986. MLPs unlimited. *Barron's* 64 (November 17): 14–20.

Terando, W. D., and T. C. Omer. 1992. Corporate characteristics associated with master limited partnership formation. *Journal of the American Taxation Association* (Forthcoming).

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (May): 817–38.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 668–680

The Association Between Nonearnings Disclosures by Small Firms and Positive Abnormal Returns

Lawrence D. Brown
SUNY at Buffalo
Kwon-Jung Kim
Chung-Ang University (Korea)

SYNOPSIS AND INTRODUCTION: We formulate and test the hypothesis that nonearnings disclosures of small, but not large, firms generally are "good news." Nonearnings disclosures are defined as disclosures by managers and outsiders about news other than earnings (e.g., stock splits, takeovers, new orders). "Good news" is defined as a positive stock price reaction at the time of the information disclosure. Our hypothesis is motivated by two lines of prior research. First, managers have incentives to disclose their private information voluntarily when they expect the effects of the information on firm value to exceed the disclosure costs (Verrecchia 1983). Second, the "firm-size differential information hypothesis," advanced by Atiase (1980, 1985) and the corroborating empirical evidence of Atiase (1985, 1987), Freeman (1987), and Bhushan (1989) suggest that incentives for information production and dissemination by outsiders are an increasing function of firm size. Thus, assuming that nonearnings disclosures concerning small firms are initiated primarily by managers, whereas those of large firms are not, small (but not large) firms' nonearnings disclosures are more likely to be good rather than bad news. Using firm-specific nonearnings disclosures, identified from the *Dow Jones News Retrieval*

We appreciate the comments of the workshop participants at Michigan State University, Rutgers University, SUNY at Buffalo, Tulane University, University of Waterloo, Washington University, and two anonymous reviewers. We are grateful to Rick Antle, Ray Ball, Vic Bernard, Dosoung Choi, Doug Hanna, Prem Jain, Tom Lechner, Jevons Lee, Gerry Lobo, Vic Pastena, Mort Pincus, Gord Richardson, Mike Rozeff, Kevin Sachs, Chuck Trzcinka, and Joe Weintrop for specific comments, and to Wonsun Paek for programming assistance. The second author gratefully acknowledges the financial support of the Deloitte and Touche Foundation.

Submitted February 1992.
Accepted November 1992.

Service data base over the 1982 to 1987 period, we show that small firms' nonearnings disclosures, on average, are associated with significant stock price increases, whereas large firms' nonearnings disclosures, on average, are valuation-neutral.

Given these results and the evidence that nonearnings disclosures are often made around the time of earnings announcements (Hoskin et al. 1986; Thompson et al. 1987), we reexamine the puzzling result of Chari et al. (1988) that on-time earnings announcements of small, but not large, firms are associated with positive abnormal returns, *unconditional* upon the nature of the earnings news. We hypothesize that this phenomenon is attributable to nonearnings disclosures of good news around the time of small firms' earnings announcements. We show that small and large firms' "pure" on-time earnings announcements are *not* associated with positive abnormal returns, and that small (but not large) firms' "contaminated" on-time earnings announcements are associated with positive abnormal returns. We conclude tht the Chari et al. (1988) results do not pertain to small firms' on-time earnings announcements *per se*, but to those that are accompanied by nonearnings news.

Key Words: *Small firms, Earnings announcements, Nonearnings disclosures, Good news.*

Data Availability: *Data used in this study may be obtained from the authors upon request.*

THE article proceeds as follows. We develop hypotheses in section I, discuss data and methodology in section II, present the empirical results in section III, then conclude with a summary and implications for future research in section IV.

I. Information Disclosure and Firm Size

Earnings disclosures are mandatory, but most firm-specific nonearnings disclosures are discretionary. Verrecchia (1983) models a manager's decision to disclose or withhold information in the presence of traders with rational expectations. In this model, managers seeking to maximize firm value have incentives to disclose private information when they expect the effects of the information on firm value to exceed the disclosure costs.[1] Thus, when managers initiate firm-specific nonearnings disclosures, these are likely to be good rather than bad news.

Managers generally do not initiate all firm-specific nonearnings disclosures. Outsiders, such as security analysts and the financial press, search out and disclose firm-specific information (Thompson et al. 1987). According to the "firm-size differential information hypothesis" of Atiase (1980, 1985), incentives for information production and dissemination by outsiders are an increasing function of firm size. Security analysts concentrate their activities on larger firms (Bhushan 1989), and the financial press publishes many more items concerning these firms (Atiase 1987; Thompson et al.

---

[1] Disclosure-related costs include costs of preparing and disseminating the information, and the potential negative effect of the information on share price. A negative effect is possible even if the information is favorable because it may be used by competitors or employees in ways harmful to the firm.

1987). With the assumption that nonearnings disclosures concerning small firms are initiated primarily by managers, these are more likely to be good rather than bad news. In contrast, with the assumption that nonearnings disclosures concerning large firms are initiated primarily by outsiders, these are likely to be valuation-neutral, on average.[2] Our first hypothesis in alternative form is:

H1: Small, but not large, firms' nonearnings disclosures generally are good-news.

Previous research indicates that nonearnings disclosures are frequently made concurrently with earnings announcements (Hoskin et al. 1986; Thompson et al. 1987). This evidence, when combined with our first hypothesis, suggests that abnormal returns for small firms' earnings announcements are likely to be affected by concurrent nonearnings disclosures of good news.

Since earnings announcements usually are reported in the financial press, small firms have the media's attention on earnings announcement dates. Small firms' managers with incentives to disclose their private information and have it disseminated will choose the earnings announcement date to do so as it is one of the few times during the year that they have the media's attention. Thus, small firms' managers are more likely to disclose nonearnings good news around earnings announcements than at other times.[3] If so, the mean security return around small firms' earnings announcements will be contaminated by nonearnings good news. Failure to recognize this confounding effect may lead researchers to conclude erroneously that small firms' earnings announcements *per se* generally are good news.

Chari et al. (1988) have documented that on-time earnings announcements of small, but not large, firms are associated with positive abnormal returns, *unconditional* upon the nature of the earnings news. Ball and Kothari (1991) attempt to distinguish between increased expected returns attributable to increasing risk during the earnings announcement period (Epstein and Turnbull 1980; Robicheck and Myers 1966) and "true" abnormal returns. After adjusting for increases in small firms' risk during the earnings announcement period, Ball and Kothari still observe positive abnormal returns.[4]

The existence of positive abnormal returns around small firms' on-time earnings announcements is puzzling because it suggests that these are *per se* good news. We hypothesize that the positive abnormal returns are *not* attributable to earnings announcements *per se*, but arise from nonearnings disclosures of good news made around this time. Our second and third hypotheses in alternative form are:

H2: Small and large firms' on-time earnings announcements unaccompanied by nonearnings disclosures are not associated with positive abnormal returns, *unconditional* upon the nature of the earnings news.

---

[2] We assume that outsiders, as a group, have equal incentives to report good and bad news. We recognize that some sell-side analysts report good news if the firms employing them underwrite securities of the firms they follow (Dugar and Nathan 1992; Lin and McNichols 1992), and that some media personnel who prefer sensationalism tend to disclose bad news. We expect that such individuals are a small part of the set of outsiders, and that their biases cancel out.

[3] "Supply of information" arguments can be supplemented by "demand for information" arguments. News services report a smaller proportion of what happens in small firms, but they will carry a story about nonearnings news when they report earnings news. Thus, it is conceivable that the media reports nonearnings disclosures that managers had released earlier, as well as previously unreported nonearnings disclosures with earnings announcements.

[4] Ball and Kothari (1991) do not segment their sample into early, on time, and late earnings announcements. Thus, their findings may not pertain to on-time announcements, the subject of this study.

H3: Small, but not large, firms' on-time earnings announcements accompanied by nonearnings disclosures are associated with positive abnormal returns, *unconditional* upon the nature of the earnings news.

II. Data and Methodology

Data

In conformance with Chari et al. (1988) and Ball and Kothari (1991), small and large firms are defined as belonging to the smallest and largest deciles, respectively, of NYSE and AMEX firms. To avoid a survivorship bias, firms are selected from the CRSP tape on the basis of market value of common equity at the beginning of each of the six years, 1982–1987. Earnings announcements and nonearnings disclosures are obtained from the *Dow Jones News Retrieval Service* (DJNRS) data base. We rely on the DJNRS for firm-specific disclosures because it covers both *The Wall Street Journal* and the *Dow Jones News Service* (Broadtape).[5]

As neither *The Wall Street Journal* nor the Broadtape subsumes the other (Thompson et al. 1987), the DJNRS reduces potential errors in identifying and the dating of firm-specific disclosures. Initially, 9,063 firm-specific disclosures are identified for the small-firm sample; 4,923 are nonearnings disclosures and 4,140 are quarterly earnings announcements. For the large-firm sample, a total of 102,627 firm-specific disclosures are identified; 97,777 are nonearnings disclosures and 4,850 are quarterly earnings annoucements.[6] These samples are used to examine whether nonearnings disclosures tend to cluster around earnings announcement dates. For tests of hypothesis H1, the samples of nonearnings disclosures for small and large firms are reduced to 4,575 and 97,454, respectively, because of unavailability of returns data in the CRSP tape. For tests of hypotheses H2 and H3, the samples of earnings announcements for small and large firms are smaller, as described below, because the hypotheses pertain to on-time earnings announcements only.

Test of Hypothesis H1

To test our first hypothesis that small, but not large, firms' nonearnings disclosures generally are good news, we examine the mean daily returns of small and large firms for two portfolios: (1) a nonearnings disclosures portfolio, *NED*, consisting of firms with at least one nonearnings disclosure on day *t*; and (2) its complement, $\overline{NED}$, a portfolio omitting firms with nonearnings disclosures on day *t*. The two portfolios are equally weighted and are formed when there is at least one nonearnings disclosure for at least one firm on a given trading day.[7] The release period for a given nonearnings disclosure is defined as the two days $(-1, 0)$, where day 0 is the date of news publication in *The Wall Street Journal*.[8] We expect *NED* to have a more positive mean daily return than

---

[5] Our maintained assumption is that *The Wall Street Journal* and the *Dow Jones News Retrieval Service* constitute the set of public disclosures regarding the firm. Nevertheless, the information may be disseminated via other sources.

[6] The DJNRS data base consists of two parts, news headline and text. The headline describes the nature of the news release, and the text contains its full description. Determination of nonearnings disclosures is based on the headline.

[7] Assume that there are n_1 firms with nonearnings disclosures on a given trading day, and n_2 firms without nonearnings disclosures on that trading day. *NED* is based on n_1 observations, and $\overline{NED}$ is based on n_2 observations.

[8] When a nonearnings disclosure appears in the Broadtape but is not subsequently published in *The Wall Street Journal*, day 0 is defined as the day after the Broadtape date.

$\overline{NED}$ for small, but not large, firms.[9] To control for possible correlation between the two portfolio returns, the difference in mean daily returns between *NED* and $\overline{NED}$ is assessed separately for small and large firms by using a *t*-test for the differenced series of daily portfolio returns.

The above test compares raw returns of the two portfolios with respect to inclusion or exclusion of firms with nonearnings disclosures. Alternatively, the significance of the average abnormal returns for nonearnings disclosures *per se* can be examined. For this purpose, we estimate abnormal returns for nonearnings disclosures using the market-adjusted method:

$$AR_{it} = R_{it} - R_{mt}, \tag{1}$$

where:

t = day relative to a nonearnings disclosure,
AR_{it} = abnormal return to security i on day t,
R_{it} = return to security i on day t, and
R_{mt} = CRSP equally-weighted market return on day t.

The two-day (−1,0) abnormal return, $AR_{i(-1,0)}$, is defined as the summation of the abnormal returns for days −1 and 0 with equation (1). The average two-day (−1,0) abnormal return, $AR_{(-1,0)}$, for a sample of nonearnings disclosures is calculated as:

$$AR_{(-1,0)} = (1/n)\sum_{i=1}^{n} AR_{i(-1,0)}. \tag{2}$$

The *t*-statistic for $AR_{(-1,0)}$ is:

$$t_{(-1,0)} = \frac{AR_{(-1,0)}}{\frac{S_{i(-1,0)}}{\sqrt{n}}}, \tag{3}$$

where $S_{i(-1,0)}$ is the standard deviation of the two-day (−1,0) abnormal returns, and n is the number of nonearnings disclosures in the sample.

Tests of Hypotheses H2 and H3

In conformance with Chari et al. (1988), a quarterly earnings announcement is considered to be on time if it is within four calendar days on either side of the previous year's quarterly earnings announcement. Thus, 1,901 and 3,879 earnings announcements are identified as on time for the small- and large-firm samples, respectively. Abnormal returns for on-time earnings announcements are estimated with equation (1). Significance of the average abnormal returns for contaminated and pure earnings announcements is assessed similarly.

---

[9] Both NED and $\overline{NED}$ include earnings announcements. The number of earnings announcements for small firms in $\overline{NED}$ is much greater than in NED. As we define the release period for a given nonearnings disclosure as the two days (−1,0), earnings announcements made on day −1, 0, or +1 will affect the two-day (−1,0) returns for nonearnings disclosures. With this three-day criterion regarding the effect of earnings announcements on NED, the number of small firms' earnings announcements in $\overline{NED}$ and NED are 3,678 and 462, respectively. If small firms' earnings announcements per se are associated with positive abnormal returns, a conservative bias is introduced into tests of our first hypothesis, making it more difficult to reject the null. In spite of this reduction in power, our first hypothesis (in null form) is shown below to be soundly rejected.

The abnormal returns estimated with the market-adjusted method do not explicitly incorporate small firms' risk increases around earnings announcements. Thus, similar to Ball and Kothari (1991), we use the capital asset pricing model in risk premium form to estimate abnormal returns for earnings announcements, controlling for potential risk shifts during the earnings announcement period. The average abnormal returns for days t through $t+\tau$ (where t is the day relative to an earnings announcement) for on-time contaminated and pure earnings announcements are the intercepts, α_c and α_p, of the cross-sectional regression in event time (Jensen's [1968] alphas):

$$R_{i(t,t+\tau)} - R_{f(t,t+\tau)} = \alpha_c D_i + \alpha_p(1-D_i) + \beta[R_{m(t,t+\tau)} - R_{f(t,t+\tau)}] + e_{i(t,t+\tau)}, \tag{4}$$

where:

$R_{i(t,t+\tau)}$ = return on security i for days t through $t+\tau$,
$R_{f(t,t+\tau)}$ = risk-free rate of return for days t through $t+\tau$, based on one-month U.S. 30-day Treasury Bill rate of returns,
$R_{m(t,t+\tau)}$ = CRSP equally weighted market return for days t through $t+\tau$,
$D_i = 1$ for contaminated earnings announcements, and 0 for pure earnings announcements,
β = systematic risk, and
$e_{i(t,t+\tau)}$ = disturbance term.

The two-day $(-1,0)$ period is typical in earnings announcement studies, but the results of Chari et al. (1988) and Ball and Kothari (1991) motivate us to include days -2 and $+1$ as well. Chari et al. find the abnormal return from day -2 to 0 to be significantly positive for small firms' earnings announcements. Ball and Kothari find the abnormal return from day -1 to $+1$ to be significantly positive for small firms' earnings announcements. If, as discussed in section I, small firms tend to disclose their nonearnings good news more often around earnings announcements than at other times, the positive abnormal returns on day -2 $(+1)$ observed by Chari et al. (Ball and Kothari) may be attributable to the effect of nonearnings disclosures of good news on these days. Thus, we focus on the four-day $(-2,+1)$ period to examine abnormal returns around earnings announcements.

We classify earnings announcements as "contaminated" if one or more nonearnings disclosures are made during the four-day trading period, and "pure" if no nonearnings disclosures are made during this trading period. Of the 1,901 on-time earnings announcements of small firms, 219 are classified as contaminated, and the remaining 1,682 as pure. For large firms, 2,311 of the 3,879 on-time earnings announcements are classified as contaminated and the remaining 1,568 as pure.

III. Empirical Results

Small, But Not Large, Firms' Nonearnings Disclosures Generally Are Good News

Table 1 presents the mean daily returns for the nonearnings disclosures portfolio and its complement for the entire time period, 1982–1987, and for each of the six years. For the small firms for the six-year period, *NED* has a mean daily return of 0.994 percent, and $\overline{NED}$ has a mean daily return of 0.121 percent. Thus, *NED* for small firms has a mean daily return more than eight times as large as $\overline{NED}$, in spite of the fact that $\overline{NED}$ contains many more earnings announcements than *NED*. A t-test of the difference

Table 1
The Effect of Nonearnings Disclosures on Mean Daily Returns of Small and Large Firms

		Mean Daily Portfolio Returns (Percent) Including/Excluding Nonearnings Disclosures[a]		
	n[b]	*NED*	$\overline{NED}$	t-*value*[c] *for* $NED-\overline{NED}=0$
Small Firms[d]				
1982	237	0.979	0.223	4.27**
1983	251	1.018	0.221	4.23**
1984	253	0.710	−0.017	2.95**
1985	250	0.926	0.125	3.74**
1986	248	1.675	0.131	3.10**
1987	248	0.662	0.050	2.26*
1982–1987	1,487	0.994	0.121	7.43**
Large Firms[e]				
1982	245	0.097	0.111	−0.59
1983	253	0.077	0.087	−0.57
1984	253	0.032	0.034	−0.15
1985	252	0.126	0.124	0.17
1986	253	0.095	0.069	1.64
1987	250	0.079	0.024	2.63**
1982–1987	1,506	0.084	0.075	1.32

[a] *NED* includes firms with one or more nonearnings disclosures on day *t*. $\overline{NED}$ excludes firms with one or more nonearnings disclosures on day *t*. *NED* ($\overline{NED}$) includes earnings announcements made with (without) nonearnings disclosures. For a given nonearnings disclosure, the release period is defined as the two days (−1,0), where day 0 is the date of news publication in *The Wall Street Journal*. The two portfolio returns are equally weighted. Nonearnings disclosures are identified from the *Dow Jones News Retrieval Service*.

[b] The number of trading days in which there is at least one nonearnings disclosure for any of the sample firms.

[c] t-values are calculated for the differenced series between the two portfolio returns.

[d] Bottom decile of NYSE and AMEX firms.

[e] Top decile of NYSE and AMEX firms.

* Significant at the 0.05 level, one-tailed test.

** Significant at the 0.01 level, one-tailed test.

in the mean daily returns of the two portfolios is significant at the 1 percent level for the entire period (*t*-value equals 7.43) and for each of the individual years (*t*-values range from a low of 2.26 in 1987 to a high of 4.27 in 1982).

For the large firms for the six-year period, *NED* has a mean daily return of 0.084 percent, and $\overline{NED}$ has a mean daily return of 0.075 percent. A *t*-test of the difference in the mean daily returns of the two portfolios is insignificant in the entire period and for five of six individual years. Moreover, in three years (1982–1984), the mean daily return of the nonearnings disclosures portfolio is actually *less* than that of its complement. These results are consistent with hypothesis H1.

Tests using market-adjusted abnormal returns, omitted from the table for simplicity, reveal similar results. The mean two-day (−1,0) abnormal return for the nonearnings disclosures of small firms for the 1982–1987 period is 1.939 percent, which is significantly greater than zero at the 1 percent level (t-value=8.41). In contrast, the mean two-day (−1,0) abnormal return for the nonearnings disclosures of large firms for the period is 0.002 percent, which is insignificantly different from zero (t-value=0.20).

Small Firms' Nonearnings Disclosures Are Made More Often Around Earnings Announcement Dates

To ascertain whether nonearnings disclosures tend to cluster around earnings announcement dates, we examine the relative proportion of total nonearnings disclosures reported around earnings announcement days relative to those made at other times. We define the earnings announcement period as the four trading days, (−2,+1), where day 0 is the earnings announcement date in *The Wall Street Journal*, and calculate the ratio of the number of nonearnings disclosures per trading day for the earnings announcement period to that of the nonearnings announcement period. We find the average of this ratio to be 1.79 and 1.09 for small and large firms, respectively, for the 1982–1987 period. In other words, small firms are 79 percent more likely to have nonearnings disclosures during earnings announcement periods than at other times, while large firms are only 9 percent more likely to do so. A test for the six-year period reveals that this difference is significant at the 1 percent level.[10] We conclude that small, but not large, firms' nonearnings disclosures cluster around earnings announcement dates.

Pure and Contaminated Earnings Announcements

Table 2 presents the test results for our second and third hypotheses; namely, that small and large firms' on-time earnings announcements unaccompanied by nonearnings disclosures are not associated with positive abnormal returns, unconditional upon the nature of the earnings news, and that small, but not large, firms' on-time earnings announcements accompanied by nonearnings disclosures are associated with positive abnormal returns, unconditional upon the nature of the earnings news.

For both small and large firms, it is evident that *pure* earnings announcements are *not* associated with significant abnormal returns during the four-day (−2,+1) period. The average abnormal returns when using the market-adjusted approach for pure earnings announcements for small and large firms are 0.387 and 0.133 percent, respectively (t-values equal 1.55 and 1.37). However, *contaminated* earnings announcements of small, but not large, firms *are* associated with significantly positive abnormal returns during the four-day (−2,+1) period. The average abnormal returns for small and large firms are 1.868 and 0.132 percent, respectively (t-values equal 2.78 and 1.47). The magnitude of the abnormal returns at the time of small firms' contaminated earnings announcements is nearly five times that of the abnormal returns at the time of their pure earnings announcements. In contrast, these two magnitudes are virtually identical for large firms. A test of significance of the difference between the contaminated and pure earnings announcements yields a significant result for small, but not large, firms (respective t-values are 2.02 and −0.01). As shown in table 2, the results are nearly identical when risk shifts are controlled for. Significant results are observed for small firms' contaminated earnings announcements and for contaminated minus pure earnings announcements, but not otherwise.

[10] The difference is found to be significant at the 1.3 percent level or better in every year but 1987.

Table 2
Average Abnormal Returns (Percent) for On-Time Contaminated and Pure Earnings Announcements of Small and Large Firms[a,b]

(*t*-values in parentheses)

	Pure	*Contaminated*	*Contaminated—Pure*
Market-Adjusted			
Small Firms	0.387	1.868	1.481
	(1.55)	(2.78)**	(2.02)*
	n=1,682	n=219	
Large Firms	0.133	0.132	−0.001
	(1.37)	(1.47)	(−0.01)
	n=1,568	n=2,311	
Jensen's Alphas[c]			
Small Firms	0.366	1.833	1.468
	(1.46)	(2.66)**	(2.00)*
	n=1,682	n=219	
Large Firms	0.132	0.119	−0.013
	(1.36)	(1.37)	(−0.10)
	n=1,568	n=2,311	

[a] An earnings announcement is on time if its date is within four calendar days of the corresponding earnings announcement in the previous year.

[b] An earnings announcement is contaminated if there are one or more nonearnings announcements during the four trading days (−2, +1) where day 0 is the date of earnings announcement in *The Wall Street Journal*. Otherwise, the earnings announcement is pure. Abnormal returns are calculated for the four days (−2, +1). Existence of nonearnings disclosures around earnings announcements is determined by using the *Dow Jones News Retrieval Service*.

[c] Jensen's (1968) alphas are the two intercepts of the cross-sectional regression in event time:

$$R_{i(-2,+1)} - R_{f(-2,+1)} = \alpha_c D_i + \alpha_p (1 - D_i) + \beta[R_{m(-2,+1)} - R_{f(-2,+1)}] + e_{i(-2,+1)},$$

where:

$R_{i(-2,+1)}$ = return on security *i* for days (−2, +1);
$R_{f(-2,+1)}$ = risk-free rate of return for days (−2, +1), based on one-month U.S. 30-day Treasury Bill rate of return;
$R_{m(-2,+1)}$ = CRSP equal-weighted market return for days (−2, +1);
$D_i = 1$ for contaminated group and 0 for pure group;
β = systematic risk; and
$e_{i(-2,+1)}$ = disturbance term.

* Significant at the 0.05 level, one-tailed test.
** Significant at the 0.01 level, one-tailed test.

In sum, the results are consistent with hypotheses H2 and H3. On-time earnings announcements unaccompanied by nonearnings disclosures are not associated with positive abnormal returns, unconditional upon the nature of the earnings news, whereas small, but not large, firms' on-time earnings announcements accompanied by nonearnings disclosures are associated with positive abnormal returns, unconditional upon the nature of the earnings news. The 11.5 percent of small firms' on-time earnings announcements that are contaminated (i.e., 219 of 1,901) appear to drive the result in Chari et al. (1988) of an association between small firms' on-time earnings announcements and positive abnormal returns.[11] Indeed, the 88.5 percent of small firms' on-time earnings announcements that are pure are not associated with significant abnormal returns.

Removing the Effect of Earnings From Abnormal Returns

An alternative explanation for the results shown in table 2 is that the contaminated subsample has a preponderance of good-news earnings information. If so, our conclusion that the positive abnormal returns associated with small firms' on-time earnings announcements, unconditional upon the nature of the earnings news, are attributable to the contaminating effect of nonearnings disclosures of good news may be invalid. To test the validity of the good-news earnings information conjecture, we estimate the regression model:

$$\begin{aligned} AR_i = {} & b_1 SMALL_i^* CONTAM_i + b_2 SMALL_i^* PURE_i + b_3 LARGE_i^* CONTAM_i \\ & + b_4 LARGE_i^* PURE_i + b_5 SMALL_i^* UEARN_i + b_6 LARGE_i^* UEARN_i + e_i, \end{aligned} \tag{5}$$

where:

AR_i = four-day $(-2,+1)$ earnings announcement period abnormal return estimated with the market-adjusted method;
$UEARN_i = E_q - E_{q-4}$, where E_q is earnings before extraordinary items and discontinued operations for quarter q, deflated by the market value of equity three days prior to the earnings announcement, purged of measurement error using the procedure of Brown et al. (1987);
$SMALL_i = 1$ if the firm is small, 0 otherwise;
$LARGE_i = 1$ if the firm is large, 0 otherwise;
$CONTAM_i = 1$ if the earnings announcement is contaminated and 0 otherwise;
$PURE_i = 1$ if the earnings announcement is pure and 0 otherwise; and
e_i = disturbance team.

We expect b_2, b_3, and b_4 to equal zero, and b_1, b_5, and b_6 to be positive. We also expect $b_1 - b_2$ to be positive and $b_3 - b_4$ to equal zero. On the basis of the findings of Foster et al. (1984) and Shevlin and Shores (1993), we expect $b_5 - b_6$ to be positive. "Earnings surprise" is defined as the error with the seasonal random walk forecast model, purged of measurement error using the procedure of Brown et al. (1987).[12] In this procedure, *UEARN* is defined as the residual in the regression: $SRW_i = c_0 + c_1 CR_i + v_i$, where *SRW* is the forecast error from the seasonal random walk model, and *CR* is stock return for days $(-252, -3)$. This procedure has been found to be effective in reducing measurement error by Easton and Zmijewski (1989) and Lipe (1990).

For the small-firm sample, 105 of the contaminated and 967 of the pure earnings announcements had years t and $t-1$ quarterly earnings data available from the *COMPUSTAT Quarterly File.* For the large-firm sample, earnings data were available for

---

[11] Similar to the findings in Chari et al. (1988), the results for the entire sample of 1,901 on-time earnings announcements for small firms are significant. The mean four-day $(-2,+1)$ abnormal return when using the market-adjusted method is 0.558 percent, and its t-value is 2.38.

[12] Other approaches, such as using analysts' or ARIMA model forecasts, are impractical. Only nine of the 219 contaminated observations for small firms are available in the *Value Line Investment Survey*. Moreover, insufficient time-series data are available to estimate an ARIMA model because the *COMPUSTAT Quarterly File* we accessed has data beginning in 1980.

2,043 of the contaminated and 1,383 of the pure earnings announcements. In an effort to increase the size of the contaminated sample of small firms, we obtained 90 additional observations from *The Wall Street Journal Index*. Observations with *UEARN*s in the top and the bottom 1 percent of this sample of 4,588 observations were excluded to avoid excessive influence by outliers on the parameter estimates.[13] The final sample consists of 4,492 earnings announcements.[14]

The results presented in table 3 conform with our expectations. More specifically, b_2, b_3, and b_4 are insignificantly different from zero; b_1, b_5, and b_6 are positive and significant; b_1-b_2 and b_5-b_6 are positive and significant; and b_3-b_4 is insignificantly different from zero.[15] Thus, tests of hypotheses H2 and H3 do not appear to be confounded by earnings information. We conclude that the positive abnormal returns around small firms' on-time earnings announcements, unconditional upon the nature of the earnings news, are attributable to good-news nonearnings disclosures rather than to good-news earnings.[16]

IV. Summary and Implications

We hypothesize and provide evidence that small, but not large, firms' nonearnings disclosures generally are good news. Motivated by these results, we extend the work of Chari et al. (1988), who show that small firms' on-time earnings announcements are associated with positive abnormal returns, unconditional upon the nature of the earnings news. We provide evidence consistent with two other hypotheses: small and large firms' on-time earnings announcements unaccompanied by nonearnings disclosures are not associated with abnormal returns, unconditional upon the nature of the earnings news; small, but not large, firms' on-time earnings announcements accompanied by nonearnings disclosures are associated with positive abnormal returns, unconditional upon the nature of the earnings news. We conclude that the finding that small firms' on-time earnings announcements are associated with positive abnormal returns is attributable to nonearnings news that accompanies small firms' on-time earnings announcements.

Our results may suggest an opportunity to reap abnormal returns by holding a small-firm portfolio around on-time earnings announcements. Similarly, one may view our results as inconsistent with market efficiency. However, it is impossible *ex ante* to determine which earnings announcements of small firms will be on time, and which of these will be accompanied by nonearnings news. Moreover, our analysis ignores transactions costs. If a 2 percent round-trip transactions cost is assumed (Ball 1992), positive returns will not exist.[17] Thus, our results should not be interpreted as substituting a nonearnings anomaly for an earnings anomaly.

---

[13] The results are insensitive to alternative methods of treating extreme observations.

[14] The sample sizes for the contaminated and pure earnings announcements are 168 and 903, respectively, for the small-firm sample, and 2,038 and 1,383, respectively, for the large-firm sample.

[15] All t-values are adjusted for heteroscedasticity by using the White (1980) procedure.

[16] Since many firms in the same industry announce their earnings on the same calendar dates, the disturbance terms in equation (5) may be positively cross-correlated, and the significance of the estimated parameters may be overstated (Bernard 1987). To alleviate this problem, we estimated equation (5) using a random sample based on industry grouping, where the random sample consists of one observation selected at random from firms in the same three-digit SIC category with the same earnings announcement dates. The results, available upon request, are similar to those reported in table 3.

[17] As shown in table 2, the average abnormal returns for on-time contaminated earnings announcements for small firms are 1.868 percent (market-adjusted method) or 1.833 percent (risk-adjusted method).

Table 3

Regression of Abnormal Returns on Unexpected Earnings for On-Time Contaminated and Pure Earnings Announcements of Small and Large Firms

(*t*-values in parentheses)

$$AR_i = b_1 SMALL_i{*}CONTAM_i + b_2 SMALL_i{*}PURE_i + b_3 LARGE_i{*}CONTAM_i + b_4 LARGE_i{*}PURE_i + b_5 SMALL_i{*}UEARN_i + b_6 LARGE_i{*}UEARN_i + e_i$$

	b_1	b_2	b_1-b_2	b_3	b_4	b_3-b_4	b_5	b_6	b_5-b_6
(Expected Sign)	(+)	(0)	(+)	(0)	(0)	(0)	(+)	(+)	(+)
	1.502 (2.21)*	−0.386 (−1.33)	1.887 (2.61)**	0.059 (0.61)	−0.023 (−0.21)	0.081 (0.58)	0.298 (5.95)**	0.153 (5.85)**	0.145 (2.56)**

Note:
AR_i = the four-day (−2, +1) earnings announcement period abnormal return, where day 0 is the date of earnings announcement in *The Wall Street Journal*;
$UEARN_i$ = the residual of the regression: $SRW_i = c_0 + c_1 CR_i + v_i$, where SRW_i is the earnings forecast error based on the seasonal random walk model, deflated by the market value of equity three days prior to the earnings announcement, and CR_i is the cumulative stock return over days (−252, −3);
$SMALL_i$ = 1 if the firm is small, 0 otherwise;
$LARGE_i$ = 1 if the firm is large, 0 otherwise;
$CONTAM_i$ = 1 if the earnings announcement is contaminated and 0 otherwise;
$PURE_i$ = 1 if the earnings announcement is pure and 0 otherwise; and
e_i = disturbance term.

All estimates except b_5 and b_6 are in percentages; $n = 4{,}492$, adjusted $R^2 = 0.037$, and $F = 29.9$, which is significant at the 0.01 level.

* Significant at the 0.05 level, one-tailed test.
** Significant at the 0.01 level, one-tailed test.

Researchers typically assume that by using a large enough sample, the effects of confounding events are randomized. This assumption is valid only if the confounding events are random. In the case of small firms' on-time earnings announcements, the confounding events are not random, but are systematic good news. Thus, the randomization assumption cannot be relied on to mitigate concerns about confounding events in this case. Future research is needed to determine if confounding effects are systematic when researchers examine other events of interest (e.g., dividend announcements).

Our finding that nonearnings news releases of small, but not large, firms generally are associated with positive abnormal returns suggests that the firm size-related return anomaly (Banz 1981; Keim 1983; Reinganum 1981) may be attributable partly to return behavior around small firms' nonearnings news releases. Although nonearnings news releases of small firms tend to be made more often around earnings announcements on a per-day basis, the total number of small firms' nonearnings news releases on days other than earnings announcement days is much greater.[18] When we examined the mean daily (raw) returns of the small-firm sample for the 1982–1987 period, the mean daily return *fell by 21.6 percent* for the period other than the first five days of January simply by excluding the approximately 3 percent of small firms per day having nonearnings news. Future research is needed to determine how much of the firm-size related return anomaly is attributable to return behavior around small firms' nonearnings news releases.

---

[18] For the 1982–1987 period, 551 nonearnings news releases were made during the four days (−2, +1) around earnings announcements, and 4,372 were made on days other than the four-day (−2, +1) period.

References

Atiase, R. 1980. Predisclosure informational asymmetries, firm capitalization, financial reports, and security price behavior. Ph.D. dissertation, University of California, Berkeley.

——. 1985. Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research* 23 (Spring): 21–36.

——. 1987. Market implications of predisclosure information: Size and exchange effects. *Journal of Accounting Research* 25 (Spring): 168–76.

Ball, R. 1992. The earnings-price anomaly. *Journal of Accounting & Economics* 14 (June/September): 319–45.

——, and S. Kothari. 1991. Security returns around earnings announcements. *The Accounting Review* 66 (October): 718–38.

Banz, R. 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics* 9 (March): 3–18.

Bernard, V. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* 25 (Spring): 1–48.

Bhushan, R. 1989. Firm characteristics and analyst following. *Journal of Accounting & Economics* 11 (July): 255–74.

Brown, L., P. Griffin, R. Hagerman, and M. Zmijewski. 1987. An evaluation of alternative proxies for the market's assessment of unexpected earnings. *Journal of Accounting & Economics* 9 (July): 159–93.

Chari, V., R. Jagannathan, and A. Ofer. 1988. Seasonalities in security returns: The case of earnings announcements. *Journal of Financial Economics* 21 (May): 101–21.

Dugar, A., and S. Nathan. 1992. The effect of investment banking relationships on financial analysts' earnings forecasts and investment recommendations. Working paper, Michigan State University, East Lansing, MI.

Easton, P., and M. Zmijewski. 1989. Cross-sectional variation in the stock market response to accounting earningsaccouncements. *Journal of Accounting & Economics* 11 (July): 117–41.

Epstein, L., and S. Turnbull. 1980. Capital asset prices and temporal resolution of uncertainty. *The Journal of Finance* 35 (June): 627–43.

Foster, G., C. Olsen, and T. Shevlin. 1984. Earnings releases, anomalies, and the behavior of security returns. *The Accounting Review* 59 (October): 574–603.

Freeman, R. 1987. The association between accounting earnings and security returns for large and small firms. *Journal of Accounting & Economics* 9 (July): 195–228.

Hoskin, R., J. Hughes, and W. Ricks. 1986. Evidence on the incremental information content of additional firm disclosures made concurrently with earnings. *Journal of Accounting Research* 24 (Supplement): 1–32.

Jensen, M. 1968. The performance of mutual funds in the period 1945–1964. *The Journal of Finance* 23 (May): 389–416.

Keim, D. 1983. Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics* 13 (June): 12–32.

Lin, H., and M. McNichols. 1992. Underwriting relationships and analysts' earnings forecasts. Working paper, Stanford University, Stanford, CA.

Lipe, R. 1990. The relation between stock returns and accounting earnings given alternative information. *The Accounting Review* 65 (March): 49–71.

Reinganum, M. 1981. Misspecification of capital asset pricing: Empirical anomalies based on earnings yield and market values. *Journal of Financial Economics* 12 (March): 19–46.

Robicheck, A., and S. Myers. 1966. Conceptual problems in the use of risk-adjusted discount rates. *The Journal of Finance* 21 (December): 727–30.

Shevlin, T., and D. Shores. 1993. Firm size, security returns, and unexpected earnings: The anomalous signed-size effect. *Contemporary Accounting Research* (forthcoming).

Thompson, R., C. Olsen, and J. Dietrich. 1987. Attributes of news about firms: An analysis of firm-specific news reported in *The Wall Street Journal Index*. *Journal of Accounting Research* 25 (Autumn): 245–73.

Verrecchia, R. 1983. Discretionary disclosure. *Journal of Accounting & Economics* 5 (December): 179–94.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (May): 817–38.

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 681–691

The Effect of Risk Factors on Auditors' Configural Information Processing

Mario J. Maletta
Boston University
Thomas Kida
University of Massachusetts at Amherst

SYNOPSIS AND INTRODUCTION: Recent audit studies by Brown and Solomon (1990, 1991) reveal that careful consideration of domain-specific knowledge can result in the experimental detection of configural relationships between information cues. These findings suggest that auditors' decisions may be more complex than indicated by some previous research, and that additional research is necessary to identify conditions in which auditors utilize configural processes.

This study investigates the role of environmental risk factors on the configurality of audit decisions. That is, we test whether the systematic consideration of audit risk variables results in the identification of higher order, interactive decision processes where linear relationships have previously been detected. The findings of Libby et al. (1985) and existing audit pronouncements are used to develop hypotheses regarding auditors' internal audit decisions.[1] The hypotheses concern (1) the interactive effect of inherent risk and control strength on the extent to which auditors rely on internal audit functions to reduce planned audit work and (2) the extent to which these environmental factors affect consideration of three components of internal audit quality: objectivity, competence, and work performed. Audit managers from a Big Six accounting firm responded to a series of audit-planning cases concerning the receivables cycle of a medium-sized manufacturing firm. Inherent risk and strength of control architecture

---

[1] Internal audit reliance refers here to auditors' decisions to reduce planned audit work because of reliance on the prior work of the internal audit function, and not to their use of internal auditors as assistants in performing the audit examination (see SAS No. 65, AICPA 1991; Maletta 1993).

We gratefully acknowledge the helpful comments of Brenda Anderson, Arnie Wright, M. Abdolmohammadi, Richard Asebrook, H. Swaminathan, Jim Smith, Jeffrey Cohen, Kathryn Sullivan, and participants in the New England Behavioral Accounting Research Series.

Submitted October 1990 and August 1991.
Accepted December 1992.

were manipulated as between-subject variables, while the objectivity, competence, and work of the internal auditors were manipulated within subjects.[2]

The results indicate that specific audit decisions are quite complex when elements of the risk environment are explicitly considered. Specifically, reliance on the internal audit function was based on a configural relationship between the levels of inherent risk and control strength. Auditors relied more on internal auditors when control architecture was strong rather than weak in conditions of high inherent risk. However, the effect of control architecture was mitigated when inherent risk was low. Also, complex relationships existed between environmental risk and task-specific components of internal audit quality. For example, auditors considered all three internal audit components when making reliance decisions in the high inherent risk and strong control strength condition, but not in certain other risk conditions. Implications of the experimental results are discussed.

Key Words: *Configural decision processing, Audit risk, Inherent risk, Control risk, Internal audit reliance.*

Data Availability: *Requests for the data may be made to the first author.*

THE remainder of the article is divided into four sections. In the first section, we address the relevant literature and develop the hypotheses. In the second, we discuss the method used. In the final two sections, we present the results and discuss the findings and implications.

I. Background and Development of Hypotheses

Brown and Solomon (1990, 1991) report evidence of configural processing by auditors when domain-specific knowledge is carefully considered and experimental materials are designed to detect such relationships. They identified configural relations in control risk evaluation tasks involving the existence or absence of compensating and amplifying controls, and in tasks involving the assessment of account misstatement risk. Their findings suggest that auditors consider configural relationships between task-specific information cues and that configural processing is predictable, given auditing knowledge specific to the task.

The studies by Brown and Solomon concentrated on audit cues that are specific to a particular task. Factors from the general audit environment, however, may also affect the configurality of auditors' decision processes. In fact, judgments and decisions that appear simple when examined in isolation may actually be quite complex when factors from the general audit environment are explicitly considered. This research investigates the effect of environmental factors on the configurality of audit decisions, specifically, auditors' decisions concerning reliance on internal auditors.

Reliance decisions were selected because previous relevant research has focused primarily on components specific to the internal audit function (i.e., objectivity, competence, and work performed) and has found that such decisions can be captured with linear, additive decision rules (see, e.g., Abdel-khalik et al. 1983; Brown 1983; Margheim 1986; Schneider 1984, 1985). These studies did not manipulate general risk factors in the audit environment, and no significant evidence of configural processing was reported. However, the findings of Libby et al. (1985), along with existing audit pronouncements, suggest that the systematic consideration of environmental risk factors will result in configural processing.

---

[2] The term "control architecture" refers to the design of the accounting control policies and procedures within the area audited (see Solomon and Shields 1993). Reference to the strength of the control architecture is effectively synonymous with the term "control design strength" as used by Libby et al. (1985).

Since previous internal audit studies that report linear decision rules required auditors to determine planned audit hours, a decision context concerning audit hours was used in the present experiment. Further, prior studies that indicate auditor configural processing (Brown and Solomon 1990, 1991) focused on evaluative judgments (e.g., assessments of the likelihood of material error). The present study therefore extends the research on configural processing by focusing on an audit planning decision context.[3]

The Role of Risk in Audit Planning

Libby et al. (1985) investigated the effects of changes in audit risk factors on auditors' general planning decisions, specifically, changes in control risk were examined within various levels of inherent risk.[4] Consistent with existing auditing pronouncements (SAS No. 55, AICPA 1988), Libby et al. described control risk as consisting of (1) control design strength, (2) test strength, and (3) test results.[5] On the basis of the general relationships described in the audit risk model (SAS No. 47, AICPA 1983), they hypothesized that the effect of changes in these elements of control risk on the level of audit work would depend on the level of inherent risk. The greater the level of inherent risk, the greater the effect a given change in control risk is likely to have on the extent of planned audit work. Thus, a decrease in control risk when inherent risk is high should have a more dramatic effect on audit risk, and thus the level of audit work, than a similar decrease when inherent risk is low.

Consistent with their stated hypotheses, Libby et al. (1985) found that the effect of control risk on the extent of planned audit work depends on the level of inherent risk. Increases in control architecture strength and test strength had a greater relative effect on decreasing planned audit work when inherent risk was high rather than low. These findings, in conjunction with existing audit pronouncements, can be used to identify the potential effects of risk factors on the configurality of auditors' decisions regarding internal audit reliance.

Control Architecture, Inherent Risk, and Internal Audit Reliance

Internal auditors monitor control systems to determine if the actual policies and procedures are performed as designed (Institute of Internal Auditors 1983). Therefore, the existence of an internal audit function can reduce the level of control risk within a particular audit area. As a result, auditors' assessments of control risk are likely to be not only a function of the strength of the control architecture, test strength, and test results, but also a function of the existence of an internal audit function. This notion is supported by audit pronouncements that describe the internal audit function as a separate, higher level of control in place to determine whether the control system is functioning effectively, and instruct the independent auditor to consider the internal audit function when evaluating the system of internal accounting control (SAS No. 9, AICPA 1975 and now SAS No. 65, AICPA 1991, and SAS No. 55, AICPA 1988).

---

[3] See Libby (1981) for a discussion of the distinction between judgment and decision making.

[4] Inherent risk is the susceptibility of an account balance to a material error, irrespective of the system of internal control. Control risk is the risk that the system of controls will not detect or prevent a material error.

[5] Control design strength was defined as the strength of the prescribed control procedures and is essentially the strength of the control architecture. Test strength is the strength of the auditors' tests of controls, and test results are effectively the outcomes of those tests.

The extent to which the internal audit function can affect auditors' assessments of control risk should depend on the control architecture monitored by the internal auditors. When the internal auditors monitor weak control architectures, their contribution toward reducing control risk should be perceived as limited. However, when they monitor compliance with relatively strong control architectures, their efforts should make a significant contribution toward reducing control risk.[6] Essentially, the extent to which the internal audit function reduces control risk should be greater when the control architecture is strong versus weak.[7]

According to the findings of Libby et al. (1985), the extent to which a change in control risk affects planned audit work depends on the level of inherent risk. A given change in control risk should result in a greater change in audit work when inherent risk is high rather than low. Thus, the extent to which the existence of an internal audit function affects auditors' planning decisions is likely to be contingent on both control architecture and inherent risk. Differences in the strength of the control architecture should have a greater effect on the extent to which the internal audit function affects planned audit work when inherent risk is high versus low.

The relationships described above also indicate that internal auditors should have the greatest relative effect on planned audit work when inherent risk is high and control architecture is strong. Specifically, the existence of an internal audit function is likely to reduce planned audit work more when the control architecture is strong rather than weak. At the same time, external auditors should be more sensitive to changes in control risk when inherent risk is high rather than low. In summary, the following hypotheses will be examined.:

H1: Control architecture and inherent risk will affect auditors' decisions regarding internal audit reliance such that (1) auditors will rely more on the internal audit function to reduce planned audit work when the control architecture is strong rather than weak, and (2) the effect of the control architecture (on such decisions) will be greater when inherent risk is high rather than low.

H2: The greatest reduction in planned audit work because of reliance on the internal audit function will occur when the control architecture is strong and inherent risk is high.

---

[6] It is possible that the internal auditors can redesign weak controls to reduce control risk in future periods. However, the existence of weak controls should limit their ability to reduce control risk in the current period.

[7] Test strength and test results were not examined in this study, given that the focus is on auditors' initial planning judgments, specifically, their reliance on the prior work of the internal auditors. Test strength and results would affect auditor judgments after compliance tests are performed, and would therefore affect the decision to perform additional audit testing. The internal auditors can also perform detection procedures and thus serve as a compensating control mechanism (see Brown and Solomon 1990), but this study is designed exclusively to address the role of the internal audit function as a monitor of the control architecture.

Control Architecture, Inherent Risk, and Internal Audit Quality

This section addresses whether combinations of inherent risk and control strength affect external auditors' consideration of internal audit quality when making reliance decisions. Previous research has indicated that these decisions are generally dependent on the quality of the internal audit function (Abdel-khalik et al. 1983; Brown 1983; Clark et al. 1981; Gibbs and Schroeder 1980; Margheim 1986; Schneider 1984, 1985). However, the extent to which differences in internal audit quality affect these decisions may be contingent upon the strength of the control architecture and the level of inherent risk. Given that the effect of the internal audit function on planned audit work is dependent on the inherent risk/control strength combination, the importance of internal audit quality should also be contingent upon such factors. Therefore, the following hypothesis will be examined:

> H3: The effect of internal audit quality on the extent to which auditors rely on internal auditors to reduce planned audit work will depend on the strength of the control architecture and the level of inherent risk.

II. Method

Audit managers from a Big Six accounting firm participated in the study while attending firm training sessions. In total, 62 usable responses were obtained and analyzed. Inherent risk and control architecture were manipulated as between-subject variables, and internal audit quality was manipulated by using three within-subject factors: internal audit competence, objectivity, and work performed. Subjects were asked to complete an audit-planning case concerning the receivables area of a medium-sized manufacturing firm. The case materials included instructions, a general description of the company and its business environment, detailed financial information, including the composition and age of the company's receivables, and a detailed description of the company's control over credit approval.

After receiving this information, subjects were given descriptions of eight different internal audit functions and one that described the company as having no internal audit department. The eight descriptions represented all possible combinations of the three internal audit variables, each manipulated at two levels. The order of the three internal audit variables was randomized across the case descriptions, and the order of the cases was randomized across subjects. For each case, the subjects were asked to determine the total number of audit hours they would be willing to budget for the purpose of auditing the collectibility of the company's receivables.[8] All of the case materials were pilot-tested with Big Six audit managers and partners to ensure that the information was clear and the task realistic.

Between-Subject Treatment Variables

SAS No. 47 suggests that sensitivity to external forces and characteristics of the industry in which a company operates should be considered when assessing inherent risk. Therefore, inherent risk was manipulated by changing the nature of the company's industry. In the high-risk condition, the company was a manufacturer and distributor of equipment and supplies to companies in the oil extraction and exploration industry. The company in the low-risk condition was a manufacturer and distributor of equipment and supplies to large medical institutions.

---

[8] The total number of audit hours budgeted represents the sum of the hours needed for tests of controls and substantive tests.

The strength of the control architecture was manipulated by changing the comprehensiveness of the company's credit approval process. In the strong-control condition, the process was described as very rigorous. For new customers, the credit department contacted the customers' vendors, bank, and the Dun and Bradstreet credit information service. For repeat customers, payment histories were evaluated extensively. In the weak-control condition, the credit approval process was much less comprehensive. The process did not involve the acquisition of extensive credit information and did not require frequent credit reviews. Pretests indicated that the manipulations for both inherent risk and control architecture were considered appropriate by audit managers.[9]

Within-Subject Treatment Variables

Internal audit quality was manipulated by changing the levels of the three within-subject treatment variables: competence, objectivity, and work performed. The manipulations of these factors were also pretested.[10]

Consistent with Schneider (1984), internal audit competence was manipulated by changing the experience level of the internal audit staff and the quality of supervision. When competence was high, the staff consisted of former Big Six auditors with at least three years of experience with Big Six firms. Also, most of the internal auditors had been with the company for at least four years, and the department head had been there for eight years. When competence was low, staff members had limited audit experience. The department consisted of individuals without college degrees as well as recent college graduates with degrees in accounting but no previous work experience.

Based on official audit pronouncements, Schneider (1984, 1985), and Abdel-khalik et al. (1983), internal audit objectivity was manipulated by changing the organizational level to which the internal audit department reported. With high objectivity, the department reported to an independent audit committee; with low objectivity it reported to the company controller.

The comprehensiveness of the compliance procedures performed by the internal auditors was used to operationalize the work-performed variable. In the high condition, the internal auditors monitored the control system by performing detailed tests to determine whether the required credit information was obtained for all customers, and whether credit decisions were consistent with company policy. In the low condition, the auditors only periodically observed credit approval activities and made inquiries of credit department personnel. In both conditions, the internal auditors reported that all credit authorization policies and procedures were operating as designed.

---

[9] Thirteen managers were asked to rate the level of inherent risk for the two firm descriptions on a scale ranging from 0 for no inherent risk to 10 for a very high level of inherent risk. The results indicated that the high and low manipulations of inherent risk were significantly different at the $p < 0.01$ level (mean responses were 9.25 vs. 4.00). The strength of the control process was rated similarly (from 0 for very weak control to 10 for very strong control). The auditors viewed the strong and weak manipulations as signficantly different at the $p < 0.01$ level (mean responses were 8.50 vs. 3.57).

[10] As before, pretest managers were given descriptions of the high and low manipulations and rated the items on 11-point scales. For example, they rated the competence of the internal audit function from 0 for a very low level of competence to 10 for a very high level. The results indicated that the high and low manipulations for each variable were significantly different at the $p < 0.01$ level (9.0 vs. 1.8 for competence, 8.6 vs. 1.8 for objectivity, and 8.9 vs. 3.6 for work performed).

Dependent Variable: Percentage Difference Scores

Percentage difference scores were used as the dependent variable. These were computed by subtracting the number of budgeted audit hours each auditor assigned in the case with no internal audit function from the number assigned in each of the eight cases describing an internal audit function. This represents the reduction in planned audit work due to reliance on the work of the internal auditors. The amount was then divided by the number of hours the auditor assigned in the case with no internal audit function. The result represents the percentage difference in budgeted audit hours specifically attributable to the existence of the internal audit function considered.

Percentage difference scores were used as the dependent variable for several reasons. Libby et al. (1985) found that inherent risk and control strength affect auditors' planning decisions regardless of whether an internal audit function is present or not. Thus, it was necesssary to distinguish between the extent to which these factors affect auditors' general planning decisions versus decisions to rely on the internal audit function. The difference scores were divided by the number of hours assigned in the case with no internal audit function to obtain a relative measure for the effect of the internal audit function. For example, if the existence of an internal audit function reduces planned audit work by five hours, it is necessary to know whether this is five out of 200 assigned hours or five out of ten hours.

III. Results

An overall analysis was initially performed with the percentage difference scores as the dependent variable, inherent risk and control strength as between-subject variables, and internal audit objectivity, competence, and work performed as within-subject variables. The data was analyzed by using analysis of variance for split plot designs.[11]

Descriptive statistics for the percentage difference scores are presented in table 1. As shown, all of the scores are negative, which indicates a decrease in planned audit hours because of the existence of an internal audit function. The results of the overall analysis, presented in table 2, indicate that the main effect for control strength is significant at the $p=0.028$ level and all three within-subject factors are highly significant ($p \leq 0.001$). In addition, the existence of several higher order interactions indicate configural cue usage. For example, the four-way interaction between inherent risk, control strength, objectivity, and work performed is significant at $p=0.018$, and the four-way interaction between inherent risk, control strength, objectivity, and competence is marginally significant at the $p=0.089$ level.

The Effect of Inherent Risk and Control Architecture

Hypothesis H1 states that auditors will rely more on the internal audit function to reduce planned audit work when control architecture is strong rather than weak, and that this effect will be greater when inherent risk is high rather than low. Consistent with the first part of this hypothesis, the significance of the main effect for control

---

[11] Homogeneity of variance tests (Cochran's) on the percentage difference scores reveal no violations of this ANOVA assumption.

Table 1
Means of Percentage Changes in Assigned Audit Hours Due to the Existence of an Internal Audit Function

(S.D. in parentheses)

Description of Internal Audit Function	*High Inherent Risk*		*Low Inherent Risk*	
	Strong Control (n=16)	*Weak Control* (n=16)	*Strong Control* (n=14)	*Weak Control* (n=16)
HO+HC+HWP	−28.9 (21.5)	−10.2 (14.8)	−19.5 (29.4)	−21.3 (22.0)
HO+HC+LWP	−20.9 (17.2)	−7.6 (12.6)	−10.8 (19.4)	−9.8 (15.8)
HO+LC+HWP	−20.3 (16.9)	−5.4 (12.7)	−7.8 (21.0)	−13.6 (19.4)
LO+HC+HWP	−20.1 (17.7)	−6.5 (12.5)	−19.4 (20.8)	−9.2 (15.3)
HO+LC+LWP	−11.3 (14.5)	−4.0 (10.0)	−9.7 (11.7)	−2.6 (6.7)
LO+HC+LWP	−14.4 (17.4)	−5.1 (10.5)	−14.8 (15.7)	−6.2 (15.1)
LO+LC+HWP	−11.9 (13.9)	−4.0 (10.0)	−9.2 (11.7)	−7.6 (17.2)
LO+LC+LWP	−10.9 (12.1)	−1.7 (8.4)	−8.3 (12.9)	−4.2 (10.6)
Between-Group Means	−17.3	−5.5	−12.4	−9.3

Note: The eight internal audit descriptions represent all possible combinations of the three internal audit variables—objectivity (*O*), competence (*C*), and work performed (*WP*)—when each is manipulated at two levels, high (*H*) and low (*L*).

Table 2
Analysis of Variance for Percentage Changes in Assigned Audit Hours Because of the Existence of an Internal Audit Function

Source of Variation	F-*Value*	p-*Value*
Main Effects		
Inherent Risk (IR)	0.03	0.867
Control Strength (CS)	5.06	0.028
Objectivity (O)	12.58	0.001
Competence (C)	18.78	0.001
Work (W)	21.79	0.001
Significant Interactions*		
O×W	7.28	0.009
IR×CS×O	7.50	0.008
IR×CS×W	4.39	0.040
IR×CS×O×C	2.99	0.089
IR×CS×O×W	5.98	0.018

* Only interactions significant at the $p < 0.10$ level are reported.

strength reported in table 2 indicates that, in general, the existence of an internal audit function affects auditors' decisions to a greater extent when control architecture is strong versus weak (the decrease in mean audit hours across inherent risk conditions was 14.8 percent with strong control vs. 7.4 percent with weak). To investigate the second part of H1, it was necessary to evaluate the effect of control strength within levels of inherent risk. *F*-tests indicate that control strength is significant when inherent risk is high ($F=7.48$, $p=0.010$). Although the means are in the appropriate direction in the low inherent risk condition, the difference is not significant ($F=0.38$, $p=0.544$). Thus, these results indicate that, consistent with the relationship hypothesized in hypothesis H1, the effect of control architecture is contingent upon the level of inherent risk. Specifically, auditors generally rely more on internal auditors when control architecture is strong rather than weak, and control architecture has a greater effect on auditors' internal audit reliance decisions when inherent risk is high rather than low.

According to hypothesis H2, the greatest reduction in planned audit work because of reliance on the internal audit function will occur when inherent risk is high and control architecture is strong. An examination of the between-group means in table 1 indicates that, consistent with the hypothesis, the greatest percentage decrease in audit hours occurred in the condition of high inherent risk and strong control (17.3 percent overall reduction vs. 5.5 percent, 12.4 percent, and 9.3 percent in the other treatment conditions). Comparisons between treatment conditions indicate that the mean reduction in the condition of high inherent risk and strong control is significantly greater than that with high inherent risk and weak control ($F=7.48$, $p=0.010$) and marginally greater than that with low inherent risk and weak control ($F=2.96$, $p=0.096$). However, the difference between high and low inherent risk with strong control did not achieve statistical significance ($F=0.86$, $p=0.361$). Thus, H2 is not fully supported, as the results indicate that the consideration of environmental risk factors results in even more complex processing than expected.

Consideration of Internal Audit Quality

Hypothesis H3 states that the extent to which auditors consider internal audit quality will be a function of inherent risk and control architecture. Consistent with this hypothesis, the significance of the four-way interaction between (1) inherent risk, control strength, internal audit objectivity, and work performed and (2) inherent risk, control strength, internal audit objectivity, and competence indicates that the effect of the internal audit quality variables depends on the levels of inherent risk and control strength. To evaluate the conditions in which the quality factors affect internal audit reliance decisions, the significance of each factor in each of the four conditions of inherent risk and control strength was analyzed.

Consistent with hypothesis H3, the results in table 3 reveal that objectivity, competence, and work performed are all significant at the $p<0.05$ level in the condition of high inherent risk and strong control ($F=9.48$, 6.98, and 7.42, respectively); none of these factors are significant in the high inherent risk and weak control condition. In addition, only competence is significant when inherent risk is low and controls are strong ($F=5.92$), while objectivity and work performed are significant with low inherent risk and weak controls ($F=6.72$ and 11.02). Note that internal audit functions were expected to have the greatest potential to affect planned audit work in the condition of high inherent risk and strong control, and that all three components of internal audit quality

Table 3
The Effects of Objectivity, Competence, and Work Performed in Between-Subject Groups

	High Inherent Risk		*Low Inherent Risk*	
Variables	*Strong Control* (*F-value*)	*Weak Control* (*F-value*)	*Strong Control* (*F-value*)	*Weak Control* (*F-value*)
Objectivity	9.48*	1.56	0.27	6.72*
Competence	6.98*	1.58	5.92*	2.74
Work Performed	7.42*	0.77	2.01	11.02*

* Significant at $p<0.05$ level.

are significant in this condition. These results indicate that the consideration of environmental risk variables increases the complexity of auditors' decisions. Consistent with hypothesis H3, inherent risk and control architecture affect the extent to which specific characteristics of the internal audit function influence auditors' reliance decisions.

IV. Discussion

The findings of this study indicate that the consideration of environmental risk factors results in the identification of complex, configural decision processing. Specifically, the study indicates that the extent to which internal audit functions affect auditors' planning decisions is not dependent solely on the simple effects of factors specifically related to the internal audit function. Instead, external auditors rely on internal auditors based on a configural relationship between inherent risk and control strength. Also, inherent risk and control strength affect the extent to which internal audit quality affects auditors' internal audit reliance decisions. These results are consistent with the major thesis of the research concerning increased complexity with the consideration of environmental risk factors. Thus, the results extend the work of Brown and Solomon (1990, 1991) and suggest that future research needs to investigate whether environmental factors such as those related to audit risk have generalizable effects on other specific decision tasks. Further, the study generalizes the results on auditor configural processing from judgment contexts to an audit decision-making context. Finally, this study also extends the work of Libby et al. (1985), by showing that risk factors not only affect auditors' general planning decisions, but also affect other specific elements of the planning process, such as auditors' internal audit reliance decisions. With respect to such decisions, the results suggest the need to investigate why auditors are not affected by internal audit quality in certain conditions of inherent risk and control strength.

As with most experimental studies certain limitations exist. Although the case materials were designed with the assistance of practicing auditors to ensure realism, in an actual audit engagement the auditor must often assimilate a broader array of information from a very complex environment. Further, the subjects were asked to make decisions on only one part of the overall audit process, and participants of only one accounting firm were used as subjects, which limits the generalizability of the results.

In summary, the results underscore the importance of considering general environmental risk factors as well as task-specific variables when examining the complexity of auditors' decision processes. Decisions that appear relatively simple may in reality be quite complex when factors from the audit environment, such as audit risk, are explicity taken into consideration.

References

Abdel-khalik, A. R., D. A. Snowball, and J. H. Wragge. 1983. The effects of certain internal audit variables on the planning of external audit programs. *The Accounting Review* 58 (April): 215–27.

American Institute of Certified Public Accountants. 1975. *Statement on Auditing Standards No. 9: The Effect of an Internal Audit Function on the Scope of the Independent Auditor's Examination*. New York: AICPA.

———. 1983. *Statement on Auditing Standards No. 47: Audit Risk and Materiality in Conducting an Audit*. New York: AICPA.

———. 1988. *Statement on Auditing Standards No. 55: Consideration of the Internal Control Structure in a Financial Statement Audit*. New York: AICPA.

———. 1991. *Statement on Auditing Standards No. 65: The Auditors' Consideration of the Internal Audit Function in an Audit of Financial Statements*. New York: AICPA.

Brown, P. R. 1983. Independent auditor judgment in the evaluation of internal audit functions. *Journal of Accounting Research* 21 (Autumn): 444–55.

Brown, C., and I. Solomon. 1990. Auditor configural information processing in control risk assessment. *Auditing: A Journal of Practice and Theory* 9 (Fall): 17–38.

———, and ———. 1991. Configural information processing in auditing: The role of domain-specific knowledge. *The Accounting Review* 66 (January): 100–19.

Clark, M. W., T. E. Gibbs, and R. G. Schroeder. 1981. How CPAs evaluate internal auditors. *CPA Journal* 51 (July): 10–15.

Gibbs, T. E., and R. G. Schroeder. 1980. External auditor criteria for evaluating internal audit departments. *The Internal Auditor* 37 (December): 34–42.

Institute of Internal Auditors. 1983. *Standards for the Professional Practice of Internal Auditing*. Altamonte Springs, FL: IIA.

Libby, R. 1981. *Accounting and Human Information Processing*. Englewood Cliffs, NJ: Prentice-Hall.

———, J. T. Artman, J. J. Willingham. 1985. Process susceptibility, control risk, and audit planning. *The Accounting Review* 60 (April): 212–30.

Maletta, M. J. 1993. Auditors' decisions to use internal auditors as assistants: The effect of inherent risk. *Contemporary Accounting Research*. (Forthcoming).

Margheim, L. L. 1986. Further evidence on external auditors' reliance on internal auditors. *Journal of Accounting Research* 24 (Spring): 194–205.

Schneider, A. 1984. Modeling external auditors' evaluations of internal auditing. *Journal of Accounting Research* 22 (Autumn): 657–78.

———. 1985. The reliance of external auditors on the internal audit function. *Journal of Accounting Research* 23 (Autumn): 911–19.

Solomon, I., and M. Shields. 1993. Judgment and decision research in auditing. In *Judgment and Decision Making Research in Accounting and Auditing*, edited by R. H. Ashton and A. H. Ashton (Forthcoming).

THE ACCOUNTING REVIEW
Vol. 68, No. 3
July 1993
pp. 692–699

BOOK REVIEWS

Shane R. Moriarity, Editor

Editor's Note: Effective August 1, 1993, two copies of books for review should be sent to Dr. James R. Boatsman, Office of the Chief Accountant, United States Securities and Exchange Commission, Washington, D.C. 20549. The policy of the *Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

JACK R. FAY, *Accounting Certification, Educational & Reciprocity Requirements: An International Guide* (Westport, CT: Quorum Books, 1992, pp. xii, 301, $65.00).

The stated purpose of this book is to provide information on accounting certification and educational requirements in various countries throughout the world. This is a complex subject and a well-written book is much needed. The book provides current information for 42 countries that responded to a questionnaire on accounting certification and education requirements. The text also discusses practitioner journals, standard-setting agencies, and, in selected countries, their code of ethics. The author provides an appropriate note of caution about using this information. He indicates the world is rapidly changing and some of the information may already be outdated. For example, the discussion of changes occurring in the EC supports this cautionary note.

The early chapters are difficult to read. For example, the chapter on reciprocity would be improved with more tables and less discussion. In this chapter, the discussion on Canadian reciprocity in ten accrediting organizations was presented in one sentence. A table would have been more effective. The tables that are included are not of good quality. Information in the tables is grouped by geographic areas which hinders comparison by dividing the information throughout the chapter. If information were also presented by similarities, it would improve the communication of the data. Additional conditions for practicing in a particular country are helpful, and the conclusion on reciprocity provides some interesting points on how to solve this issue.

Information on certification requirements might be easier to read if a list of professional certifications was provided at the beginning of each geographic zone. The discussion of certification requirements could then be limited to those unique to each country. In addition, the conclusion for this chapter would be better as an introduction, since it sets the tone for the general discussion of the information presented in the chapter.

The chapter, *Activities and Responsibilities of Professional Accountants*, presents information on differences in services and required financial statements for each country. This information is important material for international education courses. Other information important to researchers and educators concerns the influence of tax laws on GAAP, the relation of each country's policies to the IASC, and indicates those countries affected by EEC Directives. The characteristics of companies required to file financial statements is also cited.

Information on continuous education requirements is the first easy material to read. This may be due to CPE being a distinct area. Since many countries do not require CPE, these countries could have been presented in a table without comment.

The chapter, *Accounting Organizations and Journals*, is the most effective presentation in the book. Practitioner journals and lists of organizations that establish standards and influence accounting professional are capably prepared.

The chapter on ethics is easy reading. The author states that since many countries have similar codes of ethics, only a selected sample is presented. However, the book should present the other countries with a statement about which country's codes they are most closely related to. There is a good discussion of the basic code of ethics requirements and an extensive discussion for a limited number of countries.

The author prepared an excellent discussion of changes in the former Soviet Union and the changes needed in their accounting systems. The problem of AICPA reciprocity and recommended changes is also discussed. Since there are 54 licensing organizations connected with the AICPA, this has been a major problem for CPAs practicing in other countries.

Appendix A presents a list of professional accounting organizations and standard-setting bodies and provides international researchers and teachers with a comprehensive list where one can acquire information.

This book is a good addition as a supplemental text to the international accounting literature. The book provides information for anyone interested in international accounting research. Although experienced teachers with a large international literature collection may not find this research effort valuable, those seeking contacts for research, or who wish a good overview of international accounting differences, will find it an important addition.

Overall, the book provides important information for both accounting researchers and educators. It provides a good source of comparative information on many topics, as well as sources for extended research or additional information. Hopefully, shortcomings in readability and presentation will be overcome in future editions.

PAUL E. BAYES
Professor of Accountancy
East Tennessee State University

CHERYL R. LEHMAN, *Accounting's Changing Role in Social Conflict* (New York: Markus Wiener Publishing, Inc., 1992, pp. vii, 174, $34.95).

This book contains an examination of the role of accounting in certain ongoing financial crises. In doing so, it rejects the two most commonly held positions regarding that role. The position of market theorists that accounting "doesn't matter" and the position of theorists that support the view of "representational faithfulness" are both rejected. In their place the author takes a view she calls "dialectic" which assumes that accounting "both reflects and affects" struggles and is "neither an insignificant player nor a mere bearer of the facts" (3). Under this view it is important to understand both the nature of social conflict in society and the role accounting plays in the social struggle. These ideas are the focus of the analysis reported in this book.

The author utilizes a socio-historical approach in which she reviews the traditional definitions of accounting that appear in the literature and critiques the neo-classical economic model which underlies much conventional accounting research. The author rejects as illusionary such concepts as "value free, objective, and fact" as they are applied to the discipline and to accounting research. She analyzes the nature of society in which accounting and auditing practice and research have evolved and discusses the role of accounting in managing conflict through ideology. Having developed the theoretical underpinnings, the author then presents the results of semiotic classification (content analysis) of 1100 articles which appeared in the *Journal of Accountancy, Accounting Review*, and *Fortune* magazine between 1960 and 1973. The articles are classified into "discourse categories" each of which has an ideological theme. The author's conclusions are intended to provide insight into how the state managed social conflict during that period of time as well as insight into the role of accounting and accounting research.

This book is ambitious, covering a wide range of material including economic theory, philosophy of science, structure of societies, feminist theory, critical theory, and semiotics. This book is provocative in that it challenges the reader to recognize and question the most basic implicit and explicit assumptions of accounting professionalism and conventional accounting research. This book focuses on the interrelationships and interdependencies of accounting practice and accounting research in an innovative manner. The conclusions challenge the reader to view accountants as something other than impartial recorders of fact and to assess accounting's partisanship in situations of social conflict.

Accounting researchers should read this book. The issues raised are critically important to the profession of accounting. Even if one does not agree with the author, and many conventional accounting researchers will not agree, there is a benefit to be derived from reading this carefully reasoned, extensively researched book. Because it challenges the assumptions of conventional accounting research, the book would be excellent for use in Ph.D. seminars to introduce an alternative world view and alternative methodologies. Further, persons interested in the role of accounting in society would benefit from reading this thought-provoking analysis which challenges the image of the profession as an objective, neutral observer.

CRISTI K. LINDBLOM
Assistant Professor of Accounting
Babson College

KJELL ERIK RUDESTAN and RAE R. NEWTON, *Surviving Your Dissertation: A Comprehensive Guide to Content and Process* (Newbury Park, CA: Sage Publications, Inc., 1992, pp. xi, 221, $16.95, paper).

This book is written for the social science student—masters or doctoral—facing a thesis requirement. The book is divided into three parts: part I (three chapters) presents an orientation to research in the social sciences and offers some suggestions on how to develop an appropriate research topic, part II (four chapters) presents a description of the content of a social science dissertation on a chapter by chapter basis, and part III presents a potpourri of topics related to surviving a dissertation (overcoming barriers, use of a personal computer, sources of writing difficulties, etc.).

Would this book be useful to doctoral students in accounting? My answer is a qualified yes. The answer is yes because some of the chapters provide unique and valuable insights into problems that unfortunately arise in many dissertations. Students who read and absorb the content of these chapters will be able to construct better proposals and dissertations with less wasted time and effort—both their own time and effort and that of their advisors. The qualification to my recommendation arises because a considerable portion of the book is at too low a level for dissertation stage doctoral students. The text is specifically directed to an audience of masters and doctoral students and this means that doctoral students will find some of the material basic or even remedial (e.g., discussions of how to refer to the content of tables).

Students looking for a magic formula that will easily produce a suitable dissertation topic will undoubtedly be disappointed by the suggestions offered in this text. Reasonable suggestions are offered but they are not unique and they involve creative thought and effort rather than magic. In my opinion, two chapters in the book will prove especially valuable to doctoral students—the chapter concerned with literature reviews and the chapter concerned with writing in general. The authors criticize the all-too-common tendency toward thoughtless and overly long descriptive literature reviews. In addition, they provide an excellent discussion of the ordered, analytic, and purposeful nature that such reviews need to take. The chapter dealing with writing difficulties and how to overcome them argues that students' prior experiences encourage and reward writing that simply describes the ideas of others. The writing chapter specifically warns students that this aspect of their lives is about to change—their writing will have to reflect their own ideas about the meaning of the work of others and their ability to extend or develop that meaning. The literature review chapter provides a very specific application of these ideas. These two chapters would make excellent reading for all new doctoral students, and, in my opinion, these two chapters, by themselves, are enough to justify the acquisition of the book.

This book contains many ideas and recommendations about the research process. It is likely that different readers will disagree with a few of these. For example, I think that the discussion of research aimed at "merely" solving practical problems does not fit well with the current research environment in accounting. On the whole, however, the authors provide straightforward and sound advice that is not conveniently available elsewhere. I believe doctoral students at most institutions will be amply rewarded by investing a few hours in this book.

GERALD L. SALAMON
Professor of Accounting

Indiana University at Bloomington

WANDA A. WALLACE and S. DUANE SMITH, *Accounting for Options: No Longer an Option* (Vancouver: The Canadian Certified General Accountants' Research Foundation, 1991, pp. x, 192, C$19.75, paper).

This research study, funded by the Canadian Certified General Accountants' Research Foundation, examines the timely issue of accounting for employee stock options in the United States and seeks to answer if a change in the accounting method for stock options would have economic consequences. The study's components include a discussion of the standard setting process for stock options, an analysis of the market reaction to the standard setting process, a review of the valuation of stock options under alternative scenarios, an analysis of lobbying activity to the Financial Accounting Standards Board (FASB), and a review of the relevant literature.

Some of the results of the study include evidence that the standard setting activity of the FASB regarding stock options was associated with a negative stock market reaction for the sample companies. Also, the results show that other methods used and some proposed by the FASB to value options would produce a significant reduction in net income. In addition, the study provides evidence that companies with large dollar values of outstanding stock options under each of the valuation approaches illustrated by the authors are larger, have lower debt-to-equity ratios and reported income, better relative performance in the market, a larger total market value, a lower percentage of outside directors, larger boards of directors, and are more likely to lobby the FASB than firms with small dollar values of outstanding options.

The study is well written and organized and the research is well executed. The monograph is a comprehensive treatment of a complex topic. The manuscript has important implications for the FASB in its deliberations and for researchers who do work in this area. I would encourage the authors to develop articles and extensions of the study discussing these important results for wider dissemination and for different audiences.

My only criticism of the study is the authors' conclusion and opinion that current accounting rules should be revised to require a charge to income for compensation expenses related to executive stock options. Because of the results of the study which find economic consequences associated with accounting for stock options, my conclusion would be that we do not need another FASB pronouncement dealing with stock options. My feeling is that the path being proposed by the Securities and Exchange Commission of greater disclosure of compensation expense and the value of stock options in the proxy statements to be the better choice.

The study is very useful for people doing research in the executive compensation or economic consequences area. One strong feature of the book is a very comprehensive bibliography of compensation related research. Also, the book would be very good supplementary reading for those teaching a graduate seminar in financial or managerial accounting, agency theory, or executive compensation.

JAMES F. WAEGELEIN
Associate Professor

University of Kansas

PAULINE WEETMAN, editor, *SSAP 15: Accounting for Deferred Taxation* (Edinburgh: The Institute of Chartered Accountants of Scotland, 1992, pp. vi, 177, £15.00, paper).

This monograph is the sixth in the "Emerging Issues" series prepared by the Research Committee of The Institute of Chartered Accountants of Scotland. This series is addressed to those interested in accounting and financial management. This monograph reports the results of analysis of the experiences and problems resulting from accounting for deferred taxes in accordance with Statement of Standard Accounting Practice 15.

This work is consensed into 177 pages including references, bibliography, and appendixes; thus, much detail is sacrificed. Unfortunately, the brevity leaves the reader unsatisfied, and seeking other sources of information. Chapter 3, for example, discusses the present United Kingdom tax regime and the creation of timing differences in a mere seven pages. This brevity is disconcerting if the reader is not familiar with the accounting and tax system in the U.K.

There are three particulars related to this monograph that potential readers must recognize. First, non-British readers must deal with the use of unfamiliar acronyms. Second, readers will need to translate the occasional British colloquialism such as "crystallise." Third, this monograph is written like a doctoral dissertation.

In my view, the most interesting part of this monograph is the seventh chapter. It discusses the practical problems encountered in applying SSAP 15. Evidence on the practical problems was gathered from the invited comments of leading accounting firms. Each problem was enunciated and then was followed by a comment from the committee conducting the study. After discussing the problems, the committee presented both short-term and long-term solutions. At this point, all options are presented with the committee refraining from advocating one option over another. The committee's recommendations and justifications are presented in the final chapter (chapter 9) of the work.

Most readers will not benefit greatly from this work. While this monograph achieves its objectives of providing an overview of accounting for deferred taxes, this monograph is far too brief to be considered a major authoritative source.

ROBERT M. KOZUB
Associate Professor of Accounting
University of Wisconsin-Milwaukee

PAULINE WEETMAN and SIDNEY GRAY, editors, *SSAP 19: Accounting for Investment Properties* (Edinburgh: The Institute of Chartered Accountants of Scotland, 1992, pp. vi, 163, £15.00, paper).

This book is the fifth in an "Emerging Issues" series initiated by the Research Committee of the Institute of Chartered Accountants of Scotland. It is in response to a request by the Accounting Standards Board to review the experiences and problems associated with SSAP 19 *Accounting for Investment Properties* (issued November 1981). Two approaches were used in performing the research, i.e., interviews with experts and a literature review.

A series of thirteen interviews were held with a total of 22 persons of both experience and seniority in their respective areas. Out of these interviews, the Working Party gained considerable insight into current opinion and developed several accounting issues directly affecting the reported earnings of an enterprise and not adequately dealt with in SSAP 19 at present. Among the problems identified were the (1) potential for switching between group occupancy and third party occupancy, (2) potential for switching between development and investment property, (3) lack of clarity regarding temporary and permanent diminutions of value, (4) calculating and reporting gains and losses on disposal, and (5) revaluation gains and losses in associated undertakings and joint ventures. Areas for potential improvement of disclosure were also identified, e.g., more explanation about the assumptions of valuation, particularly the "reasonable period" involved and the assumed yield, and more information about the nature of the investment property portfolio.

A fairly comprehensive literature survey was also conducted. While there was relatively little directly addressing SSAP 19, there was a lot of interest in a range of issues relevant to it, both domestically and internationally, such as valuation of assets on a basis other than historical cost. One conclusion of the literature review, as with the interviews, was that while there was not a great deal wrong with SSAP 19, there still was a rich variety of issues to explore.

Based on this research, the Working Party made a number of recommendations to the Accounting Standards Board concerning the accounting for investment properties addressing the issues raised by the interviews and the literature review. For example, the lack of depreciation of investment property was not found to be an issue at the practical level, though conceptually the academic literature review did raise some concerns about it which ought to be addressed in SSAP 19. Also, among others, the Working Party recommended that a number of the issues raised by the experts during the interview process be dealt with more clearly and decisively.

In summary, the book provides a useful and clearly written introduction and analysis of important issues concerning accounting for investment properties which may be of interest in accounting classes discussing depreciation and valuation of different types of assets, or to accountants, regulators, and others who are interested in asset valuation and disclosure problems.

JOHN T. RIGSBY
Associate Professor
Mississippi State University

STEPHEN A. ZEFF, FRANZ VAN DER WEL, and KEES CAMFFERMAN, *Company Financial Reporting: A Historical and Comparative Study of the Dutch Regulatory Process* (Amsterdam: North-Holland, 1992, pp. xv, 410, $85.00).

In the initial chapter of this book, the authors point out that three industrial enterprises with substantial or exclusive Dutch participation are of commanding international importance, and each have shares listed on the New York Stock Exchange. The three multinational enterprises are the Royal Dutch/Shell Group, the biggest enterprise in Europe and second largest in the world with 60 percent ownership by the Dutch holding company, Phillips, the eighth largest in Europe; and Unilever with 50 percent ownership by the Dutch holding company and the twelfth largest in Europe. The accountants and auditors for these three companies as well as with several other large Dutch firms have had considerable influence on the development of accounting and auditing in the Netherlands. Further, the Amsterdam Stock Exchange, which now ranks eighth among stock exchanges in Europe, is the world's oldest stock exchange. Such information along with a sketch of the Dutch national government and its political parties, the legal structure of company organizations, the auditing profession, and university accounting education provide the setting for the historical study which is the core of this book. Key comparative developmental factors in this setting are the irrelevance of the tax law, the absence of a securities commission, and the lack of an accounting "standards" setting program.

The historical analysis of Dutch accounting practice is divided into five developmental periods, namely, the early company law development from the late 19th century until 1945, but going back as far as the 1837 Commercial Code; the agitation for improved financial reporting from 1945 to 1959; the swift political current toward a new company law from 1959 to 1970; the decade of the tripartite accounting study group in the 1970s; and the first decade of the Council on Annual Reporting from 1981 to 1991. A chapter is devoted to each of these historical periods.

Though the regulatory process itself is very important and a key development in this book, its effects on certain financial reporting issues such as secret reserves, segment reporting, and consolidated financial statements are also quite important to the accounting academician. Significant attention is given to these emerging issues in the context of financial reporting developments. Because the book has a subject index and a name index, the historical tracing of any of these key reporting issues can be readily accomplished. Actually, some of these concepts are now taken for granted in international financial reporting. How they developed so recently in the Netherlands, quite frequently through peer pressure from international reporting circles, are among the most interesting topics in the book.

One of the hurdles in undertaking an English language book where much of the background research is in a foreign language is to translate key terms and nomenclatures into English in a manner that maintains the original flavor. Throughout the book, key Dutch terms and names are given in both English and Dutch. Also, though acronyms are used generously, their meanings are repeated sufficiently often so that the flow of the reading is smooth. Only one rather important Dutch term, "bedrifseconomie," is not translated. Since the term envelopes a combination of managerial economics and business administration, the authors decided not to give it a translation. In this regard, some concepts such as "true information" (p. 64) and "accounting elegance" (p. 77) as early financial reporting objectives might be found amusing to the American reader.

The book is the result of a project undertaken at the joint invitation of the Limperg Institute and the Dutch Council on Annual Reporting to provide a record of the historical development of the efforts toward improved financial reporting in the Netherlands, and to make recommendations for possible reform. In other words, this book is the first and only comprehensive historical treatise on the Dutch accounting regulatory process. Very important to the book's development is the availability to the authors of the confidential minutes and papers of the key study groups and annual reporting councils and committees that have developed financial reporting recommendations. Further, the authors interviewed key participants and observers in the regulatory process, including auditors, company executives and other representatives that were active on these study groups, councils, and committees. These sources provide a uniquely thorough historical study, documenting the efforts of key dedicated professionals that had more than a national vision for improved financial reporting in their country.

Intriguing and quite valuable to the reader are the country comparisons which follow the historical study. The comparative analysis of national cultures, capital markets, legal systems, and regulatory environments between the Netherlands, the United States, and the United Kingdom provides further significant insights into the understanding of the Dutch regulatory process. This cross-cultural analysis furnishes the backdrop for the concluding chapter with its policy recommendations.

As a comprehensive historical study and comparative analysis of the development of the accounting regulatory process in the Netherlands, the book is a major contribution to the accounting literature. For international accounting and historical researchers and educators, the book will be a vital reference work.

NORLIN G. RUESCHHOFF
Professor of Accountancy

University of Notre Dame

BOOKS RECEIVED NOVEMBER 1, 1992 TO FEBRUARY 5, 1993

DAVID GRAYSON ALLEN and KATHLEEN McDERMOTT, *Accounting for Success: A History of Price Waterhouse in America 1890–1990* (Boston: Harvard Business School Press, 1993, pp. xx, 373, $35.00).

JOHN ARNOLD, DON EGGINTON, LINDA KIRKHAM, RICHARD MACVE, and KEN PEASNELL, *Goodwill and Other Intangibles* (London: The Institute of Chartered Accountants in England and Wales, 1992, pp. xii, 93, £10.00, paper).

G. JACK BOLOGNA, ROBERT J. LINDQUIST, and JOSEPH T. WELLS, *The Accountant's Handbook of Fraud and Commercial Crime* (New York: John Wiley & Sons, Inc., 1993, pp. xi, 339, $105.00).

ROBERT J. CAMPBELL, *Competitive Cost-Based Pricing Systems for Modern Manufacturing* (Westport, CT: Quorum Books, 1992, pp. 259, $55.00).

ROBIN COOPER, ROBERT S. KAPLAN, LAWRENCE S. MAISEL, EILEEN MORRISSEY, and RONALD M. OEHM, *Implementing Activity-Based Cost Management: Moving from Analysis to Action* (Montvale, NJ: Institute of Management Accountants, 1992, pp. xviii, 336, $35.00).

DOUGLAS D. DAVIS and CHARLES A. HOLT, *Experimental Economics* (Princeton: Princeton University Press, 1993, pp. xiii, 572, $39.50).

PATRICK R. DELANEY, JAMES R. ADLER, BARRY J. EPSTEIN, and MICHAEL F. FORAN, *GAAP Interpretation and Application 1993 Edition* (New York: John Wiley & Sons, 1992, pp. xxxiv, 785, $42.00 paper).

JEFFREY L. FUNK, *The Teamwork Advantage: An Inside Look at Japanese Product and Technology Development* (Norwalk, CT: Productivity Press, 1992, pp. xxxi, 477, $50.00).

DAVID I. GOLDENBERG, *The U.S. Man-Made Fiber Industry: Its Structure and Organization Since 1948* (Westport, CT: Praeger Publishers, 1992, pp. xii, 281, $49.95).

PELHAM GORE, *The FASB Conceptual Framework Project 1973–1985: An Analysis* (Manchester, UK: Manchester University Press, 1992, pp. xi, 1989, $69.95).

CHRISTOPHER HUMPHREY, PETER MOIZER, and STUART TURLEY, *The Audit Expectations Gap in the United Kingdom* (London: the Institute of Chartered Accountants in England and Wales, 1992, pp. viii, 99, £10.00, paper).

YUJI IJIRI and ISAO NAKANO, editors, *Business Behavior and Information* (Pittsburgh: Carnegie Mellon University Press, 1992, pp. xvi, 200, $15.00, paper).

G. H. LAWSON, editor, *Studies in Cash Flow Accounting and Analysis: Aspects of the Interface Between Managerial Planning, Reporting and Control and External Performance Measurement* (New York: Garland Publishing, Inc., 1992, pp. vi, 315, $63.00).

DAVID L. McKEE and DON E. GARNER, *Accounting Services, The International Economy, and Third World Development* (Westport, CT: Praeger Publishers, 1992, pp. xi, 161, $47.95).

JOSEPH M. MORRIS, *Software Industry Accounting* (New York: John Wiley & Sons, Inc., 1993, pp. xviii, 299, $95.00).

MARILYN KLEINBERG NEIMARK, *The Hidden Dimensions of Annual Reports: Sixty Years of Social Conflict at General Motors* (New York: Markus Weiner Publishing, Inc., 1992, pp. v, 207, $34.95).

DONALD J. PETERS, *A Contrarian Strategy for Growth Stock Investing: Theoretical Foundations & Empirical Evidence* (Westport, CT: Quorum Books, 1993, pp. xiv, 177, $49.95).

PAULINE WEETMAN, editor, *SSAP 15 Accounting for Deferred Taxation* (Edinburgh: The Institute of Chartered Accountants of Scotland, 1992, pp. vi, 177, £15.00, paper).

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal.... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of the *Chicago Manual of Style* (13th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's International Dictionary*.

FORMAT

1. All manuscripts should be typed on one side of 8½ × 11" good quality paper and be double spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To assure anonymous review, authors should not identify themselves directly or indirectly in their papers. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title, and affiliation, any acknowledgements, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: *All* pages, including tables, appendices, and references, should be serially numbered. The first section of the paper should be untitled and unnumbered. Major sections may be numbered in roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights, and measures. For example: *three days*; *3 kilometers*; *30 years*. All other numbers are expressed numerically. Generally when using approximate terms spell out the number, for example, *approximately thirty years*.

Percentages and Decimal Fractions: In nontechnical copy use the word *percent* in the text; in technical copy the symbol % is used. (See the *Chicago Manual* for discussion of these usages.)

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: a *well-presented analysis*; *re-form*. See *Webster's* for correct usage.

Key Words: The synopsis is to be followed by four *key words* that will assist in indexing the paper.

SYNOPSIS

A synopsis of about 400 words should be presented on a separate page immediately preceding the text. The synopsis also serves as the "introduction." The synopsis should be nonmathematical and include a readable summary of the research question, method, and the significance of the findings and contribution. The title, but not the author's name or other identification designations, should appear on the synopsis page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an arabic number and a complete title indicating the exact contents of the table or figure.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix *a*, *b*, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says. . . ."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by the *Chicago Manual:*

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Dates of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) should be listed in chronological order of publication. Two or more works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2d ed. Homewood, IL: Dow Jones-Irwin.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, United Kingdom: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–36.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, University of Texas, Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be double spaced and numbered consecutively throughout the manuscript with superscript Arabic numerals. Footnotes are placed at the end of the text.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or other publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan, or the like) should be submitted.
3. Four copies should be submitted together with a check in U.S. funds for $50.00 for members or $100.00 for nonmembers of the AAA made payable to the American Accounting Association. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the Editor for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting.

"An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional, research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 703–724

The Market Valuation of Accounting Information: The Case of Postretirement Benefits other than Pensions

Eli Amir
Columbia University

SYNOPSIS AND INTRODUCTION: The majority of empirical regulatory accounting studies on financial reporting have focused on the *ex post* evaluation of accounting choices (see Lev [1979] and Barth [1991] among others) after an accounting method is adopted. There has not been much research involving the evaluation of accounting methods *ex ante* because, it is argued, newly mandated accounting information is usually available only some time after the Financial Accounting Standards Board (FASB) has required its release.

In December 1990, after lengthy and controversial deliberations, the FASB issued SFAS No. 106, *Employers' Accounting for Postretirement Benefits other than Pensions* (PRB) replacing the current "pay as you go" practice with a combination of present-value and accrual method. Few of the information items mandated by SFAS No. 106 has yet been disclosed earlier. Consequently, this research uses the data that were available during the time of the FASB's deliberations as an example of *ex ante* empirical research concerning the standard-setting process.

This study uses the PRB cash payments to retirees as disclosed by firms in their footnotes to the financial statements under SFAS No. 81 (FASB 1984) to investigate whether investors underestimated the full effect

This paper draws on my Ph.D. dissertation completed at University of California—Berkeley. I would like to thank Trevor Harris, Jonathan Leonard, Joshua Livnat, David Modest, Stephen Penman, Jacob Thomas, Thomas Rothenberg, Paul Ruud, anonymous referees, and seminar participants at the University of California—Berkeley, Columbia University, University of California—Davis, Duke University, Harvard University, New York University, University of British Columbia, and Washington University for helpful comments. Baruch Lev and Brett Trueman deserve special thanks for their careful review and guidance. The financial support of the Deloitte Touche Foundation is gratefully acknowledged.

Submitted January 1992.
Accepted March 1993.

Editor's Note: This article was awarded the AAA Competitive Manuscript Contest Award for 1992.

of the PRB liability on firms' values. Such underestimation would occur if investors are not aware of the high rate of increase in health care costs or when future benefits' payments to current employees are partially ignored. Additionally, this study investigates whether an *estimate* of the present value of the PRB liability is value-relevant to investors. Finally, an analysis is made of the range and sensitivity of economic parameters (discount rate and health care cost trend rate) used by investors to estimate the PRB obligation.

The results indicate that, during the period 1984–1986, investors, on average, valued each dollar of PRB cash payment in any year as a dollar; i.e., they underestimated the full consequences of firms' promise to continue making such payments in the future. During the period 1987–1990, in contrast, investors translated each dollar of PRB cash payment to an average of $13.75 PRB obligation. Further, estimating the present value of the PRB obligation with publicly available data shows that the present-value measure is value-relevant to investors in addition to the cash payments disclosed by firms.

Key Words: *Postretirement benefits other than pensions, FASB, Valuation model, Present value.*

Data Availability: *Data used in this study were obtained from public sources. A list of sample firms is available from the author upon request.*

THE remainder of this study is organized as follows: Section I provides background information about the issue and its importance. The valuation model used in this study and the tests conducted are discussed in section II. Section III describes the sample selection and the data collection process. The results of the evaluation of the value-relevance of the PRB cash payments disclosure mandated by SFAS No. 81 are presented in section IV. Section V describes the estimation method used to estimate the present value of the PRB obligation. Evidence on the value-relevance of the estimated PRB obligation is reported in section VI. Section VII contains concluding remarks.

I. Accounting for PRB—Background

Preliminaries

About 82 percent (41 percent) of U.S. companies with more (less) than 1,000 employees provide nonpension postretirement benefits, such as health care and life insurance, to former employees and their dependents (Foster Higgins 1988). The U.S. General Accounting Office estimated the aggregate *unfunded* PRB obligation of private employers at approximately $402 billion in 1988 (GAO 1989). During the period 1985–1990, the average total spending of large and mid-sized firms on retiree health care increased by 80 percent, an average annual increase of 12.4 percent. Possible reasons for this increase are a growing pool of elderly retirees, more cost-shifting by the federal government from the Medicare program to employer-provided retiree health care plans, early retirement, longer life expectancy, and court rulings that have barred employers from terminating postretirement health care plans (Geisel 1987).

In November 1984, the FASB issued SFAS No. 81, *Disclosure of Postretirement Health Care and Life Insurance Benefits*, which mandates disclosure of firms' annual cash payments to retirees for postretirement health care and life insurance benefits. In February 1989, the FASB issued an Exposure Draft, *Employers' Accounting for Postretirement Benefits other than Pensions*, which proposed to end the "pay as you go" accounting practice used by firms with PRBs, and to require that they use accrual accounting.[1]

In December 1990, the FASB approved SFAS No. 106. The FASB concluded that PRBs, like pensions, are part of employees' deferred compensation, and therefore should be recognized on the basis of service rendered, in a manner similar to the provisions of SFAS No. 87 (FASB 1985). For plans sponsored by U.S. public companies, the standard became effective for fiscal years beginning after December 15, 1992. For non-U.S. and small plans (fewer then 500 participants), the standard will become effective in fiscal year 1995.

II. Research Design

The Valuation Model

To analyze the valuation issues discussed above, this study uses a valuation model that relates the value of the firm to the information provided in the income statement (earnings) and balance sheet (book value of equity).

The valuation model derived by Ohlson (1991) is adopted here. Unlike other valuation models used in the literature, this model relates equity value to both accounting earnings and book value of equity. Both are value-relevant because they assist in predicting future dividends, the theoretical basis for a firm's market price. The model can be written as follows:

$$P_t = \gamma_o \mathrm{x}_t + \gamma_1 y_t + \gamma_2 d_t + \beta \circ \mathbf{V}_t + \eta_t, \tag{1}$$

where P_t is the market value per share of the firm's common stock at time *t*; $\mathbf{x}_t$ is the firm's earnings per share over the period $(t-1, t)$; y_t is the firm's book value per share at time *t*; d_t is the firm's dividends per share at time *t*; $\mathbf{V}_t$ is a vector of other value-relevant information; γs are parameters that relate earnings, book value, and dividends to value; βs are parameters that relate other information to value; and η_t is a disturbance term.

Substituting the "clean surplus equation," $y_t = y_{t-1} + \mathrm{x}_t - d_t$, into equation (1) yields the following relation:[2]

$$P_t = \alpha_0 y_{t-1} + \alpha_1 \mathrm{x}_t + \alpha_2 y_t + \beta \circ \mathbf{V}_t + \eta_t. \tag{2}$$

By using the clean surplus equation, price is expressed solely as a function of accounting earnings, book values, and other value-relevant information.

---

[1] Espahbodi et al. (1991) document a significant negative market reaction around the announcement of the Exposure Draft for firms with a PRB plan. Amir and Livnat (1992) document a significant positive market reaction around firms' announcement of early adoption of SFAS No. 106.

[2] The "clean surplus relation" is simply the accounting identity where the change in book value is equal to earnings minus dividends.

A potential problem with the model, as reflected in both equations (1) and (2), is heteroscedasticity. To correct for it, equation (2) is deflated by the book value of equity at time $t-1$ (y_{t-1}). The following model is then used:

$$\frac{P_t}{y_{t-1}}=\alpha_0+\alpha_1\frac{x_t}{y_{t-1}}+\alpha_2\frac{y_t}{y_{t-1}}+\beta\frac{V_t}{y_{t-1}}+\eta_t^*. \tag{3}$$

Equation (3) represents the deflated version of Ohlson's model. Other information is defined as value-relevant if it is not reflected in earnings (x) or book value (y), and $\beta\neq 0$.

Analysis of the Valuation of PRB Cash Payments

This section addresses the valuation implications of the cash payments made to retirees under the PRB plan. By analyzing the valuation effects of the PRB cash payments, one can infer whether investors properly estimate the magnitude of the firm's promise to pay PRB cash payments and, to some extent, how they translate cash payments into a measure of PRB obligation. This analysis is conducted by using the following model:

$$\frac{P_{it}}{y_{it-1}}=\alpha_{0t}+\alpha_{1t}\frac{x^a_{it}}{y_{it-1}}+\alpha_{2t}\frac{y_{it}}{y_{it-1}}+\beta_{1t}\frac{c_{it}}{y_{it-1}}+\beta_{2t}\frac{pex_{it}}{y_{it-1}}+\beta_{3t}\frac{upl_{it}}{y_{it-1}}+\eta^1_{it}, \tag{4}$$

where P_{it} is firm i's per share market price at time t; x^a_{it} is firm i's earnings per share plus the per share tax-adjusted[3] PRB cash payments expensed in year t, plus the per share tax-adjusted pension expense for year t; y_{it} is firm i's book value per share at time t; $y_{i,t-1}$ is firm i's book value per share at time $t-1$; c_{it} is firm i's tax-adjusted per share amount of cash paid by firms to their *retirees* during the year (per SFAS No. 81); pex_{it} is firm i's tax-adjusted pension expense per share; upl_{it} is firm i's per share unfunded pension liability, calculated as the difference between the accumulated benefit obligation (ABO) and the fair value of pension assets. The ABO measure of the pension liability is used primarily because data is available for the whole test period. Lastly, η^1_{it} is the normally distributed error term with a zero mean. In this model, the PRB cash payments, c_{it}, and the pension variables should be thought of as the "other information" variables.

Firms with higher PRB cash payments to retirees are expected to have larger obligations. Therefore, β_1 is expected to be negative because it measures the extent to which investors capitalize current cash payments in determining firm value. β_1 can be interpreted as a multiplier that investors use each period to translate after-tax cash payments to retirees into the tax-adjusted present value of the PRB obligation. A coefficient below -1 will indicate that investors regard PRB cash payments as permanent. Note that β_1 may change over time because of changes in the economic parameters affecting the size of the obligation (i.e., discount rate and health care cost inflation) or because of changes in how investors perceive the PRB cash payments (i.e., transitory or permanent).[4] Therefore, changes in β_1 over time will yield information on how investors update their estimates of these obligations over time.

---

[3] The tax adjustment is made by estimating the *effective* tax rate for each firm in each year. The effective tax rate is defined as income tax expense divided by pre-tax income. This rate is bounded between zero and 40 percent during 1984–1986 and between zero and 35 percent during 1987–1990.

[4] In the present context, a transitory expense/payment is one that is not expected to repeat itself in the future. A permanent expense/payment is expected to repeat itself in the future.

Analysis of the PRB Obligation

As in the case of the PRB cash method, it is important to understand how a measure of the PRB obligation is related to the firm's market value. Is the obligation used in an unbiased manner by investors in valuing firms? How sensitive is the PRB obligation to the choice of different parameters used to estimate the obligation, and what set of parameters is used by investors in valuing firms? The most relevant question, however, is whether a measure of the present value of the PRB obligation, based on publicly available information, is value-relevant to investors in addition to the cash (pay as you go) information. To understand how investors use the present value of the PRB obligation in valuing the firm's equity, the following cross-sectional valuation model is estimated with 1990 data:

$$\frac{P_{i,90}}{y_{i,89}}=\alpha_0+\alpha_1\frac{x^a_{i,90}}{y_{i,89}}+\alpha_2\frac{y_{i,90}}{y_{i,89}}+\beta_2\frac{pex_{i,90}}{y_{i,89}}+\beta_3\frac{upl_{i,90}}{y_{i,89}}+\beta_4\frac{L_{i,90}}{y_{i,89}}+\eta^2_i, \tag{5}$$

where the subscript 90 (89) refers to the 1990 (1989) fiscal year-end, $L_{i,90}$ is the firm's estimated PRB liability, and all other terms are as before.

If investors estimate the tax-adjusted present value of the PRB liability in an unbiased manner, and if the estimate used here is an unbiased estimate, then β_4 would be about -1. However, since the estimated liability is sensitive to the choice of the discount rate and the health care inflation rate, β_4 will differ from -1 according to the set of parameters used. Therefore, the size of β_4 can indicate the range of parameters used by investors in estimating the liability.

The analysis is conducted in two stages. First, a measure of the PRB liability is examined to see if it is value-relevant in addition to the cash payments information. Second, the accounting numbers are adjusted such that earnings and book values include the estimated present values of the PRB obligation. The cash information is then examined to see if it is value-relevant to investors in addition to the liability information.

The first stage of this analysis uses the following model (which is models [4] plus [5]):

$$\frac{\mathrm{Pi}_{90}}{y_{i,89}}=\alpha_0+\alpha_1\frac{x^a_{i,90}}{y_{i,89}}+\alpha_2\frac{y_{i,90}}{y_{i,89}}+\beta_1\frac{c_{i,90}}{y_{i,89}}+\beta_2\frac{pex_{i,90}}{y_{i,89}}+\beta_3\frac{upl_{i,90}}{y_{i,89}}+\beta_4\frac{L_{i,90}}{y_{i,89}}+\eta^3_i. \tag{6}$$

If the estimated liability is value-relevant in addition to cash, then β_4 will be negative and significantly different from zero, which leads to the first hypothesis.

H1: The liability estimate is value-relevant in addition to cash ($\beta_4<0$).

In the second stage of the analysis, the accounting earnings, book value, and lagged book value for each firm are adjusted to include the liability. These adjustments and transformations are presented as follows:

$$y^1_{i,90}=y_{i,90}-L_{i,90}, \tag{7a}$$

$$y^1_{i,89}=y_{i,89}-\frac{L_{i,90}+c_{i,90}}{1+D}, \tag{7b}$$

$$x^1_{i,90}=x^r_{i,90}+pex_{i,90}+c_{i,90}-D*\frac{L_{i,90}+c_{i,90}}{1+D}; \tag{7c}$$

where D denotes the discount factor used to estimate the liability (discount rate and health care growth rate), and all other terms are as defined before.

The following equation is then estimated:

$$\frac{P_{i,90}}{y^1_{i,89}} = \alpha_0 + \alpha_1 \frac{x^1_{i,90}}{y^1_{i,89}} + \alpha_2 \frac{y^1_{i,90}}{y^1_{i,89}} + \beta_1 \frac{c_{i,90}}{y^1_{i,89}} + \beta_2 \frac{pex_{i,90}}{y^1_{i,89}} + \beta_3 \frac{upl_{i,90}}{y^1_{i,89}} + \eta^4_i. \qquad (8)$$

In this formulation, β_1 is not expected to be significant, since earnings and book value include sufficient PRB information. This yields the second hypothesis:

H2: Cash payments are not incrementally value-relevant to investors beyond the estimate of the present value of the liability ($\beta_1 = 0$).

III. Sample Selection and Data Collection

The 1990 COMPUSTAT data base was surveyed, and all firms that contributed to a defined-benefit pension plan were identified as potential sample firms. Firms with pension plan assets less than $25 million were deleted from the sample in order to increase the significance of postretirement benefits in the firm's operations. The initial sample contained 556 firms. The footnotes to the 1984–1990 financial statements were read, and the firms were classified into two groups according to their SFAS No. 81 disclosure: (1) firms that disclosed some PRB information, and (2) firms that did not disclose any PRB information. In 1990, 409 firms (74 percent of the initial sample) disclosed some information regarding a PRB plan, and 147 firms (26 percent of the initial sample) did not.[5]

Data Collection

Starting in December 1984, SFAS No. 81 required each firm to disclose information regarding its PRB plan in footnotes to the financial statements. This information includes the total cash payments to retirees and some descriptive items about the plan. As noted, 409 firms included some information regarding their PRB plan in 1990. However, some of these firms state that the PRB plan is insignificant with respect to the firm's operations, and did not disclose the dollar amount paid in cash to retirees. These firms were deleted from the analyses. The firms with the highest cash payments were the automobile manufacturers GM and Ford. Total PRB cash payments for the sample are $8.3 billion in 1989 and $9.2 billion in 1990, which represents about 80 percent of total PRB payments made by the U.S. private sector (FERF 1989).

The 1990 COMPUSTAT industrial file was used to calculate the remaining accounting variables: earnings per share before extraordinary items, effective tax rate, book value per share, pension expense per share, unfunded pension liability per share, industry classification, and PRB payments per share. All variables were adjusted for stock splits and stock dividends. The 1990 CRSP files were used to retrieve price data for the fiscal year-end; demographic data and data used to estimate the PRB liability are described in section V. The sample selection process for the years 1984–1990 is summarized in table 1.

---

[5] The test sample (409 firms) includes 30 (7.3 percent) public utilities, 16 (3.9 percent) financial firms, 22 (5.4 percent) retail and service firms, 34 (8.3 percent) oil and gas firms, 18 (4.4 percent) transportation firms, and 289 (70.7 percent) manufacturing firms.

Table 1
Sample Selection

Year	1990	1989	1988	1987	1986	1985	1984
Potential sample size	556	556	556	556	556	556	556
Firms with no PRB data	147	178	175	168	187	194	232
Firms with insignificant PRB cash payments	82	60	60	59	57	53	53
Firms for which other COMPUTSTAT/CRSP data is not available	9	9	15	26	19	37	24
Removed as outliers	5	5	5	5	4	6	4
Number of observations	313	304	301	298	289	266	243

Note: Firms with market-to-lagged-book ratios of above 15 and below 0.1 were removed as outliers.

IV. Valuation of PRB Cash Payments

Panel A of table 2 provides some statistics on the regression variables, and panel B presents the sample correlation matrix. Table 3 provides the results of estimating equation (4) over the period 1984–1990. As can be seen, the α_1 coefficient (earnings-to-lagged-book ratio) varies between 3.37 (1985) and 7.069 (1989), which suggests that transitory components in earnings change *significantly* over time. The coefficient on PRB cash payments over lagged book value, β_1, is negative in five of the six years presented (it is positive in 1986), and is significantly different from zero in each of 1987–1990.[6] In 1987–1990, β_1 is below -12 and significantly different from -1. This suggests that investors are discounting future cash payments because they believe them to be a permanent commitment—obligation.

The β_1 coefficient decreases (becomes more negative) almost consistently over time. When the data are pooled, there appears to be a change in investors' assessment of the PRB obligation from the period 1984–1986 to 1987–1990. An F-test to detect a structural change (Chow test) between the two periods yields an F-statistic that is significant at the 1 percent level. A two-tailed *t*-test to examine the change in β_1 from 1984–1986 to 1987–1990 indicates that the change is significant at a 1 percent level.

Dummy regressors were also used to test whether β_1 in each of the years 1987–1990 is significantly smaller than β_1 in each of the years 1984–1986. In each regression, data were pooled from one year in 1987–1990 with data from one year in 1984–1986 Intercept and slope dummies allow the regression coefficients to vary over time to test the significance of the difference in β_1s over time. The *t*-tests showed that *each* of the coefficients in the 1987–1990 period is significantly smaller than *each* of the coefficient

---

[6] When the data are pooled over the period 1987–1990, β_1 is negative and significant.

Table 2
Sample Statistics

Panel A. Descriptive Statistics, 1984–1990 (s.e. in parentheses):

	$\frac{P_{it}}{y_{it-1}}$	$\frac{x^a_{it}}{y_{it-1}}$	$\frac{y_{it}}{y_{it-1}}$	$\frac{c_{it}}{y_{it-1}}$	$\frac{pex_{it}}{y_{it-1}}$	$\frac{upl_{it}}{y_{it-1}}$
1990:						
Mean	1.79	0.13	1.06	0.009	0.010	−0.10
	(1.60)	(0.15)	(0.21)	(0.010)	(0.020)	(0.29)
1989:						
Mean	2.25	0.17	1.09	0.009	0.009	−0.18
	(1.55)	(0.15)	(0.29)	(0.014)	(0.017)	(0.25)
1988:						
Mean	1.92	0.17	1.08	0.007	0.010	−0.12
	(1.05)	(0.13)	(0.19)	(0.011)	(0.026)	(0.41)
1987:						
Mean	1.90	0.15	1.08	0.007	0.011	−0.13
	(1.10)	(0.14)	(0.17)	(0.014)	(0.034)	(0.48)
1986:						
Mean	1.85	0.11	1.03	0.007	0.012	−0.14
	(1.07)	(0.13)	(0.21)	(0.013)	(0.027)	(0.20)
1985:						
Mean	1.59	0.11	1.04	0.005	0.016	−0.09
	(0.75)	(0.13)	(0.15)	(0.006)	(0.019)	(0.18)
1984:						
Mean	1.40	0.15	1.04	0.004	0.017	−0.09
	(0.70)	(0.11)	(0.13)	(0.009)	(0.020)	(0.17)

Panel B. Full-Sample Correlation Matrix (2,014 firm/year observations):

	$\frac{P_{it}}{y_{it-1}}$	$\frac{x^a_{it}}{y_{it-1}}$	$\frac{y_{it}}{y_{it-1}}$	$\frac{c_{it}}{y_{it-1}}$	$\frac{pex_{it}}{y_{it-1}}$	$\frac{upl_{it}}{y_{it-1}}$
P_{it}/y_{it-1}	1.00	0.60	0.41	−0.03	0.00	−0.12
x^a_{it}/y_{it-1}	0.66	1.00	0.60	0.16	0.19	−0.05
y_{it}/y_{it-1}	0.44	0.65	1.00	0.05	0.01	−0.06
c_{it}/y_{it-1}	−0.09	0.08	0.03	1.00	0.50	0.35
pex_{it}/y_{it-1}	0.00	0.22	0.07	0.18	1.00	0.47
upl_{it}/y_{it-1}	−0.22	−0.16	−0.10	0.16	0.25	1.00

Note: In the matrix, numbers *above* the diagonal represent Pearson correlations, and numbers *below* the diagonal represent Spearman correlations. Sizable differences between Spearman and Pearson correlations may indicate that outliers play a significant role in the analysis. To avoid this problem, the entire analysis was repeated with truncation to eliminate outliers, with no significant effect on the results reported here. The variables are defined as follows:

P_{it}/y_{it-1} = market value of common equity per share, divided by lagged book value of equity per share,
x^a_{it}/y_{it-1} = earnings per share adjusted for PRB and pension expenses divided by lagged book value of equity per share,
y_{it}/y_{it-1} = current book value of equity per share, divided by lagged book value of equity per share,
c_{it}/y_{it-1} = PRB cash payments to retirees per share, divided by lagged book value of equity per share,
pex_{it}/y_{it-1} = pension expense per share, divided by lagged book value of equity per share, and
upl_{it}/y_{it-1} = unfunded pension liability per share, divided by lagged book value of equity per share.

Table 3
Regressions of PRB Model, 1984–1990
(p-values for a two-tail test in parentheses)

Model: $\frac{P_{it}}{y_{it-1}} = \alpha_{0t} + \alpha_{1t}\frac{x^a_{it}}{y_{it-1}} + \alpha_{2t}\frac{y_{it}}{y_{it-1}} + \beta_{1t}\frac{c_{it}}{y_{it-1}} + \beta_{2t}\frac{pex_{it}}{y_{it-1}} + \beta_{3t}\frac{upl_{it}}{y_{it-1}} + \eta^1_{it}.$

Year(s)	α_0	α_1	α_2	β_1	β_2	β_3	$\bar{R}^2$
1990	0.047	6.619	0.969	−14.580	−4.761	−0.143	0.516
($n=313$)	(0.908)	(0.000)	(0.019)	(0.000)	(0.228)	(0.596)	
1989	1.013	7.069	0.150	−12.680	−9.982	−0.576	0.454
($n=304$)	(0.001)	(0.000)	(0.610)	(0.010)	(0.047)	(0.067)	
1988	1.191	4.454	0.042	−12.690	8.877	−0.296	0.331
($n=301$)	(0.000)	(0.000)	(0.899)	(0.026)	(0.000)	(0.047)	
1987	2.382	5.868	−1.180	−18.250	5.534	−0.221	0.371
($n=298$)	(0.000)	(0.000)	(0.003)	(0.000)	(0.047)	(0.144)	
1986	1.163	5.061	0.115	0.551	−3.007	−0.405	0.431
($n=289$)	(0.000)	(0.000)	(0.695)	(0.921)	(0.239)	(0.113)	
1985	1.278	3.370	0.080	−3.312	2.533	−0.269	0.398
($n=266$)	(0.000)	(0.000)	(0.812)	(0.688)	(0.270)	(0.205)	
1984	1.578	4.945	−0.780	−0.412	−6.312	0.102	0.397
($n=243$)	(0.000)	(0.000)	(0.033)	(0.931)	(0.007)	(0.675)	
1987–1990	0.771	5.823	0.328	−13.752	2.293	−0.343	0.415
($n=1{,}216$)	(0.000)	(0.000)	(0.055)	(0.000)	(0.168)	(0.001)	
1984–1986	1.008	3.882	0.166	−0.772	−4.551	−0.323	0.341
($n=798$)	(0.000)	(0.000)	(0.404)	(0.982)	(0.002)	(0.026)	

clients in the 1984–1986 period (at a 5 percent level). This further supports the conclusion of a material change in investors' perception of the PRB obligation.[7]

It appears that during 1984–1986, investors underestimated the present value of current and future health care for retirees and other nonpension benefits relative to the period 1987–1990. During 1987–1990, investors translate \$1.00 of cash payment to an average of \$13.75 of obligation, with β_1 ranging from −12.68 to −18.25. This significant change in investors' perception could be explained by one or more of the following: (1) investors changed their assessment of the future increase in health care costs; (2) the discount rate (settlement rate) changed; (3) investors updated their assessment of future PRB payments (scope of the plan); (4) court cases led investors to change their beliefs about the legal status of the PRB obligation; (5) public debate about this issue drew more attention to it.

Over the period 1985–1990, the PRB multiplier (β_1) has risen from 3.31 to 14.58 in absolute values, indicating that changes in investors' assessments of the discount rate or the health care cost trend rate are unlikely to be the only reason for the change in the β_1 coefficient.[8]

---

[7] The analyses were repeated by using the Heckman (1979) procedure, which attempts to detect and correct sample selection bias. The inverse Mill's ratio was derived from a probit model that included size, industry dummies, location dummies, labor intensity, and the ratio of retirees to active employees. The results were not materially affected by the inclusion of the inverse Mill's ratio in the regression model.

Changes in the coverage of PRB plans also are unlikely to have caused such significant changes in investors' assessment of the PRB liability. Over the period 1984–1990, overall coverage had not increased and most firms were trying to contain the PRB costs. A telephone survey of more than 200 financial officers (see Bacon et al. 1990) revealed that retirees will be paying more in the future for their own health benefits.

Furthermore, in 1984, U.S. district court used the common law to mandate vesting of retiree health benefits. In 1989, the sixth U.S. Circuit Court of Appeals ruled that "if plan documents clearly and consistently reserved a right to modify a plan, then modifications, based on those documents, are not illegal" (Geisel 1989). This suggests that legal proceedings may not be responsible for investors' changing beliefs.

That investors were less aware of the consequences of paying medical expenses to retirees during the period 1984–1986, is a plausible explanation. In contrast, during the period 1989–1990, as a result of the FASB Exposure Draft, public hearings, SFAS No. 106, and numerous publications in the financial press, investors could have become more cognizant of the size of the PRB liability.[9]

The coefficient on the pension expense (β_2) variable changed magnitude and sign almost every year over the 1984–1990 period, which is consistent with the results found in Grant (1991) and Barth et al. (1992). When the data were pooled, β_2 is negative and significant in the 1984–1986 period, and positive and insignificant for the 1987–1990 period. The coefficient on the unfunded pension obligation (β_3) is more stable over time. When the data were pooled, β_3 is negative and significant over 1984–1986 and 1987–1990. This result is consistent with Landsman (1986). The analysis was repeated without the pension variables and showed that β_1 values over 1984–1990 are not affected by the inclusion of the pension expense and the unfunded pension liability as regressors.

V. Estimating the Present Value of the PRB Liability with Publicly Available Information

The method of estimating the *firm-specific* PRB liability (present value) with publicly available data uses a general present-value model. Consider a firm that sponsors a PRB plan that entitles eligible employees to full medical coverage when they retire from the firm. The employee may retire on or after age y_r; n_j denotes the number of retirees of age j sponsored by the firm; and m_i the number of active employees of age i working for the firm. Expected medical costs for a retiree of age y_r for the current year are denoted by C_{y_r}. The cost is assumed to increase with age at a rate of g_1 per year. Also, because of inflation, medical costs are assumed to increase at a rate of g_2 per year. That means that a retiree of age y_r today is expected to spend (and therefore be reimbursed by the firm) $C_{y_r}[(1+g_1)(1+g_2)]^t$ for medical costs incurred t years from today, if he or she is alive at that time. For simplicity, assume that all payments are made at the end of the year. The nominal interest rate, r, is assumed constant over time. For each individual, the probability of living through year a is denoted by $PL(a)$. The probability of staying with the firm between age a and b is denoted by $PS(a,b)$. The probability of retiring at age p (where p is greater than or equal to y_r) is $PR(p)$. Under the above assumptions, the present value of the expected obligation of the firm is given by: [10]

---

[8] To get a present-value multiplier of 3.31, assuming a 20-year horizon, the combined annual discount factor should be 30 percent. That means that if the discount rate is 9 percent, the annual health care cost trend rate is around −21 percent. To get a present value of 14.58 under the same assumptions, the combined discount factor is 3.2 percent, and the annual change in health care costs is around 5.8 percent.

[9] There appears to be a market inefficiency because the footnote information on the size of cash payments for postretirement benefits was available in the financial statements during the time when postretirement benefits were relatively undervalued by investors.

$$PV=\left[\sum_{j=y_r}^{\infty} n_j C_{y_r}\left(\sum_{k=j}^{\infty}(1+g_1)^{k-y_r}(1+g_2)^{k-j}(1+r)^{j-k}PL(k)\right)\right]$$

$$+\left\{\sum_{i=0}^{\infty} m_i PS(i,y_r)C_{y_r}(1+g_2)^{y_r-i}(1+r)^{i-y_{r\circ}}\right.$$

$$\left.\left[\sum_{q=0}^{\infty}PR(y_r+q)\left(\sum_{l=j}^{\infty}(1+g_1)^{l-y_r}(1+g_2)^{l-j}(1+r)^{j-l}PL(k)\right)\right]\right\}. \qquad (9)$$

Equation (9) has two components: The first component, which represents the expected actuarial present value of paying health benefits to retirees, takes into account the expected increase in health care costs because of aging, g_1, and inflation, g_2. The second component represents the present value of future PRB payments to employees, currently active, and takes into account the probability of their being employed by the firm until retirement age, $PS(i,y_r)$, as well as the distribution of all possible retirement ages, $y_r+q(q>0)$. Since data on age groups and on other parameters in the model are generally not publicly available, some assumptions are required to make the model estimable.

Consider now the case of an employee who becomes eligible for PRB benefits upon reaching age 55, which most firms use as the eligibility age in their PRB plan (see FERF 1989 and Foster Higgins 1988). Assume that the firm usually hires new employees when they are 35 years old, and all employees retire at age 62 (ages 35 and 62 represent average employment and retirement age according to Labor Department statistics). The choice of three age groups is made mainly because of data limitations. In group 1, n_1 retirees of age 62 are eligible for PRB coverage; in group 2, n_2 employees of age 55 are fully eligible for PRB and will retire at age 62 if they live until that age; in group 3, n_3 employees of age 35 are not eligible for PRB.

Expected medical costs for a 62-year-old retiree for the current year are denoted by C_{62}. The cost is assumed to increase with age and over time in the manner described above. For simplicity, assume that an employee does not leave the firm before retirement except through death. Under the above assumptions, the present value of the expected obligation for the whole firm, which is the sum of the obligations for the three groups, is:

$$PV=S\left[n_1+n_2\left(\frac{1+g_2}{1+r}\right)^7 PS(55,62)+n_3\left(\frac{1+g_2}{1+r}\right)^{27} PS(35,62)\right], \qquad (10)$$

---

[10] Derivation of actuarial models can be found in Ippolito (1986) and Mittelstaedt and Wharshawsky (1992).

where:

$$S=\sum_{j=1}^{\infty} C_{62}(1+g_1)^j(1+g_2)^j(1+r)^{-j}PL(62+j).$$

The date of estimation is December 31, 1990. The variables in equation (10) that require estimation are the health care cost trend rate (g_2), discount rate (r), average health costs per retiree (C_{62}), the three age groups (n_1, n_2, n_3), and the probabilities of staying employed [$PS(\bullet)$] and alive [$PL(\bullet)$].

Health Care Cost Trend Rate

To estimate the rate of increase in per capita health expenditures (*PCHE*), the following notation will be used: *GNP*=gross national product (annual), *PCGNP*=per capita gross national product, *CPI*=consumer price index, *MCCPI*=medical component of the consumer price index, *HE*=national health expenditure (annual), PCHE =per capita health expenditure, Q=quantity of goods produced in the economy, *QH*=quantity of health services produced in the economy, *n*=general population size. A solid dot (•) above a variable represents its percentage change over one year.

Taking derivatives with respect to time:[11]

$$G\dot{N}P=C\dot{P}I+\dot{Q}. \tag{11}$$

Equation (11) means that the percentage change in the GNP is equal to the percentage change in prices plus the percentage change in the quantity of goods produced (real growth). Similarly, for health services only,

$$H\dot{E}=MC\dot{C}PI+Q\dot{H}. \tag{12}$$

The per capita versions of equations (11) and (12) are:

$$PC\dot{G}NP=C\dot{P}I+\dot{Q}-\dot{n}, \tag{13}$$

$$PC\dot{H}E=MC\dot{C}PI+Q\dot{H}-\dot{n}. \tag{14}$$

Four scenarios are considered here. The first scenario (S1) is that of no real growth in per capita medical services ($Q\dot{H}=\dot{n}$), that medical costs will grow at the same rate as general inflation ($C\dot{P}I=MC\dot{C}PI$).

$$PC\dot{H}E=C\dot{P}I. \tag{15}$$

Some believe that the minimum estimate of future change in health care costs could be based on the general inflation rate. Proponents of this approach believe that the CPI is more reliable than other measures (see FERF [1989, 103] for further discussion of the issue).

With the second scenario (S2), real growth in per capita medical services ($Q\dot{H}>\dot{n}$) is allowed, but medical costs still grow at the same rate as general inflation ($C\dot{P}I=MC\dot{C}PI$).

$$PC\dot{H}E=C\dot{P}I+Q\dot{H}-\dot{n}. \tag{16}$$

---

[11] Equation (11) is derived as follows:

$$\frac{\partial GNP/\partial t}{GNP}=\frac{\partial(CPI*Q)/\partial t}{CPI*Q}=\frac{\partial(CPI)/\partial t}{CPI}+\frac{\partial Q/\partial t}{Q}=C\dot{P}I+\dot{Q}.$$

Equations (12) through (14) are derived similarly.

The third scenario (S3) places no restriction on real growth or prices of medical services.

$$PC\dot{H}E = MC\dot{C}PI + Q\dot{H} - \dot{n}. \tag{17}$$

Scenarios S1 and S2 are special cases of scenario S3. Scenarios S2 and S3 are less conservative and allow for higher forecasts of the health care cost trend rate. Because economists argue that the share of *HE* from the *GNP* cannot be higher than a certain percentage (see GAO 1989; U.S. Department of Commerce 1989),[12] a fourth scenario (S4) is added to restrict the share of *HE* not to exceed 15 percent of *GNP* in 30 years.[13] For $P\dot{C}HE$ to rise from 11.1 to 15 percent in 30 years, the annual increase must be $(15/11.1)^{1/30}$ or 1 percent. This yields the following equation:

$$P\dot{C}HE = PC\dot{G}NP + 1 = C\dot{P}I + \dot{Q} - \dot{n} + 1. \tag{18}$$

The next step is to forecast $PC\dot{H}E$ according to these four scenarios. First, the annual change in the general inflation rate $(C\dot{P}I)$ must be forecast. The pattern of inflation has changed considerably during the last 50 years (see U.S. Department of Commerce 1975, 1986, 1989). An estimate of 4.5 percent as FERF (1989) will be used.[14] Second, a forecast of the $MC\dot{C}PI$ is needed. The average *MCCPI* over the period 1950–1987 was 6 percent, and the average for 1984–1987 was 6.6 percent, which is about 2 percent over the $C\dot{P}I$. A forecast of 6.5 percent will be used here. To forecast $Q\dot{H}$, the techniques of Box and Jenkins (1976) are used to identify and estimate the following model over the period 1960–1987 (*t*-values are in parentheses):

$$Q\dot{H}_t = \underset{(3.9)}{0.496\,Q\dot{H}_{t-1}} + \underset{(3.8)}{0.493\,Q\dot{H}_{t-5}}. \tag{19}$$

With this model, $Q\dot{H}$ is forecast to be, on average, 4 percent over the next 30 years. The U.S. general population $(\dot{n})$ has been growing at an average annual rate of 1.5 percent (see U.S. Department of Health and Human Services 1988), and this is the forecast used here. As for $\dot{Q}$, I adopt the U.S. Department of Commerce (1986) forecast, which is 2 percent on average for the next 30 years.

Using the above forecasts, one can now calculate $PC\dot{H}E$ under the four scenarios.

S1: $PC\dot{H}E = C\dot{P}I = 4.5$ percent.
S2: $PC\dot{H}E = C\dot{P}I + Q\dot{H} - \dot{n} = 7.0$ percent.
S3: $PC\dot{H}E = MC\dot{C}PI + Q\dot{H} - \dot{n} = 9.0$ percent.
S4: $PC\dot{H}E = C\dot{P}I + \dot{Q} - \dot{n} + 1\% = 6.0$ percent.

The forecasts result in very different estimates of the PRB obligation. Scenarios S1 and S4 would be more acceptable to the GAO and the U.S. Department of Labor, but scenarios S2 and S3 yield higher estimates.

Discount Rate

Four possible discount rates may be appropriate for calculating the PRB obligation. The first possibility is the settlement rate (*D*1), which is the rate required to "sell" the obligation to an outside party. This may be the appropriate discount rate because it measures the marginal interest rate that the firm pays to an independent agent to meet the PRB obligation. The problem is that firms are currently unable to "sell" the PRB obligations. Therefore, the settlement rate may not be identifiable by the firm. Indeed, firms disclose their own calculations of the pension rates in their footnotes to the annual financial statements, and this rate is used here.

---

[12] For example, if $PC\dot{G}NP = 2$ percent and $MC\dot{C}PI - C\dot{P}I = 4$ percent, then the share of *HE* from the *GNP* will rise from 11.1 percent (the share of *HE* from the *GNP* in 1988) to 20.1 percent in 30 years. This is considered by economists to be extremely unlikely.

[13] GAO (1989) uses 14.9 percent and the U.S. Department of Labor (1986) uses 12.8 percent.

[14] Using time-series models may give unreliable forecasts. The average $C\dot{P}I$ over the period 1950–1987 was 4.25 percent, and the average over 1982–1987 was 3.8 percent. GAO (1989) uses an annual forecast of 5 percent.

The second possibility is the marginal borrowing rate of the firm ($D2$). Assuming the firm has to borrow funds to pay the PRB liability, the marginal borrowing rate may be the appropriate discount rate. This rate is set equal to the average yield on a bond with the same S&P rating as that of the firm. The S&P bond rating for each firm is taken from COMPUSTAT, and the average yield on each class of bonds for the year 1989 is taken from the S&P Bond Guide. These rates are: $AAA=8.70$ percent, $AA=8.99$ percent, $A=9.51$ percent, $BBB=10.01$ percent, $BB=12.07$ percent, and $B=13.85$ percent.[15]

The third possibility is the return on equity ($D3$). Employers may self-finance the PRB obligation by issuing stock. In that case, the return on equity may be the appropriate rate to use. To estimate the return on equity, the market model is estimated by using monthly returns over the period 1984–1989.

The last possibility for the discount rate is the average cost of capital of the firm ($D4$). The funds to pay future PRB claims may come from both borrowing and issuing new captial. Therefore, the appropriate rate would be the average cost of capital calculated as the weighted average of $D2$ and $D3$ as follows:

$$(D4)=\frac{BD}{BD+ME}(D2)+\frac{ME}{BD+ME}(D3), \tag{20}$$

where BD is the book value of the firm's debt, and ME is the market value of the firm's equity. Book value of debt is used because the market value of debt is usually not available, and the book value serves as a good proxy for it. The sample averages for $D1$ through $D4$ are 8.9, 10.3, 18.5, and 13.3 percent, respectively. The corresponding medians are 9, 9.5, 18.2, and 13.2 percent.[16]

Average Per Capita Health Care Costs and Demography

The average cost per retiree is calculated as:

$$\frac{\textit{cash payments for PRB}}{\textit{\# retirees}}.$$

---

[15] If the firm's bonds were unrated, the yield is estimated by using the following ratio:

$$\frac{\textit{Interest on Long-Term Debt}}{\textit{Long-Term Debt}+\textit{Current Portion of Long-Term Debt}}$$

This ratio is restricted to the interval 8.7 through 13.85 percent to be consistent with the interval AAA through BB used for rated bonds.

[16] More than 470 comment letters were received by the FASB in response to its 1989 Exposure Draft. Also, five days of public hearings on this topic were held in October and November 1989. In a comment letter, Unocal Corp. argued that the discount rate should be the firm's cost of capital. Others, including the Financial Executives Institute (FEI), claimed that a normalized bond rate should be used instead. FEI, Chrysler, and others argued that the CPI index should be used as the appropriate measure of the health care cost trend rate. The Academy of Actuaries, GAO, and IBM, among others, supported the FASB's opinion that a "best estimate" should be used. S&P Corp. argued that no estimate can be reliable enough and that the present value of the liability is too sensitive to the choice of this parameter.

Although annual cash payments to retirees are disclosed by firms in their annual financial statements, firm-specific data by age group are not generally available. This study uses the *ERISA Blue Book of Pension Funds* (1983, 1984–1985, 1985–1986), which is based on form 5500 filed by each firm with the U.S. government, as a source of information for number of firm employees and retirees for some specific years. These data may be used to estimate the number of retirees and the number of active employees age 55 and older. The following notation is used throughout this section:

R_t = the number of retirees at the end of year t;
A_t = the number of active employees at the end of year t;
$R_t(a,b)$ = the number of retirees between ages a and b at the end of year t;
$A_t(a,b)$ = the number of employees between ages a and b at the end of year t;
$PL(a,b)$ = the probability of staying alive between age a and b.

As before, assume all employees retire at age 62. Assume further that the ratio of retirees to active employees and the number of employees age 55 and over to total employees remain fixed over time. Then:

$$\frac{R_{89}}{A_{89}}=\frac{R_{84}}{A_{84}}=\frac{R_{83}}{A_{83}}=\frac{R_{82}}{A_{82}}. \tag{21}$$

$$\frac{A_{89}(55,62)}{A_{89}}=\frac{A_{84}(55,62)}{A_{84}}=\frac{A_{83}(55,62)}{A_{83}}=\frac{A_{82}(55,62)}{A_{82}}. \tag{22}$$

Given these assumptions, the number of retirees in 1989 can be presented as follows:

$$R_{89}=R_{82}\circ PL(62,69)+\sum_{k=55}^{61}A_{82}(k)\circ PL(k,k+7). \tag{23}$$

Because of data availability problems, equation (23) is approximated as follows:

$$R_{89}\approx R_{82}\circ PL(62,69)+A_{82}(55,62)\circ PL(55,62). \tag{23}$$

Equation (23) approximates the number of retirees in 1989 as the sum of two components: (1) the number of retirees in 1982 that are expected to live until 1989 and (2) the number of active employees between age 55 and 62 who are expected to live until age 62 and retire at that age. Using equations (22) and (23) and rearranging terms:

$$A_{82}(55,62)=\frac{A_{89}R_{84}}{A_{84}PL(55,62)}-R_{82}\left[\frac{PL(62,69)}{PL(55,62)}\right]. \tag{24}$$

Using equations (22) and (24):

$$A_{89}(55,62)=\frac{A_{89}}{A_{82}}\left\{\frac{A_{89}R_{84}}{A_{84}PL(55,62)}-R_{82}\left[\frac{PL(62,69)}{PL(55,62)}\right]\right\}. \tag{25}$$

Equation (25) is an estimate of the number of employees between ages 55 and 62. The number of active employees age 55 and under will be the difference between $A_{89}(55,62)$ and A_{89}. The number of retirees in 1989 is given by the left-hand side of equation (23).

Probabilities of Staying Employed and Other Parameters

This study assumes that the probability of staying with the firm depends only on the event of death. This is a reasonable assumption for employees age 55 and over, because

Table 4
Estimated PRB Liability as a Percentage of Book Value

Panel A. Sample Medians:

Discount Rate[a]	Health Inflation Rates[b] S1	S2	S3	S4
D1	0.25	0.48	0.90	0.38
D2	0.21	0.36	0.62	0.29
D3	0.06	0.08	0.10	0.08
D4	0.10	0.14	0.21	0.12

Panel B. Correlations Between Liability Over Lagged Book Value ($L_{i,90}/y_{i,89}$) and Other Variables:

Variable[c]	Correlation
P_{it}/y_{it-1}	−0.27
x^a_{it}/y_{it-1}	−0.13
y_{it}/y_{it-1}	−0.20
c_{it}/y_{it-1}	0.65
pex_{it}/y_{it-1}	0.17
upl_{it}/y_{it-1}	−0.14

[a] The discount rates are: D1 = pension; D2 = bond; D3 = equity; and D4 = cost of capital.
[b] The health inflation rates are: S1 = 4.5 percent; S2 = 7.0 percent; S3 = 9.0 percent; and S4 = 6.0 percent.
[c] The variables are defined in table 2.

employees nearing retirement age are more likely to stay with the firm. For younger employees, this may not be a good assumption, but data on employee turnover are not available. Each probability of staying employed (alive) is thus calculated with actuarial life tables published by the U.S. Department of Health and Human Services (1988).

It is also assumed that the level of health care expenditure does not change with the retiree age ($g_1 = 0$ percent in equation [10]). This assumption is made because data on health care costs for different age groups were not available.

By using all the data described above, the PRB liability was estimated for each firm according to equation (10). The estimates are calculated with each of the four discount and health inflation rate combinations. Descriptive statistics for the 231 firms in the sample without missing data are reported in table 4. The sample medians of the estimated PRB liability for each pair of health care cost trend rate and discount rate (panel A) show that the liability measures are very sensitive to the choice of parameters. The average correlation between the liability estimates and the other variables used in the analysis are shown in panel B.

VI. Valuation of the Estimated PRB Liability

As expected, the PRB liability and annual cash payments are highly correlated. Given this high correlation, does the estimate of the PRB liability add value-relevant information to investors in addition to the cash payments information? The results of

Table 5
Regressions of Cross-Sectional Valuation Model with 1990 Data
(231 observations)

$$\text{Model: } \frac{P_{i,90}}{y_{i,89}} = \alpha_0 + \alpha_1 \frac{x^{o}_{i,90}}{y_{i,89}} + \alpha_2 \frac{y_{i,90}}{y_{i,89}} + \beta_2 \frac{pex_{i,90}}{y_{i,89}} + \beta_3 \frac{upl_{i,90}}{y_{i,89}} + \beta_4 \frac{L_{i,90}}{y_{i,89}} + \eta_i^2.$$

S	D	α_0	α_1	α_2	β_2	β_3	β_4	P(−1)	$\bar{R}^2$
1	1	−0.520	7.640	1.350	−9.439	0.031	−0.369		0.60
		(0.161)	(0.000)	(0.000)	(0.049)	(0.908)	(0.094)	(0.01)	
1	2	−0.497	7.620	1.338	−9.409	0.052	−0.526		0.61
		(0.183)	(0.000)	(0.000)	(0.050)	(0.847)	(0.085)	(0.11)	
1	3	−0.414	7.781	1.267	−8.900	0.131	−1.775		0.62
		(0.266)	(0.000)	(0.000)	(0.056)	(0.627)	(0.011)	(0.84)	
1	4	−0.386	7.733	1.272	−7.610	0.141	−1.531		0.62
		(0.300)	(0.000)	(0.000)	(0.110)	(0.601)	(0.022)	(0.76)	
2	1	−0.557	7.577	1.366	−10.102	0.012	−0.134		0.59
		(0.131)	(0.000)	(0.000)	(0.034)	(0.964)	(0.221)	(0.00)	
2	2	−0.545	7.549	1.356	−10.212	0.025	−0.179		0.59
		(0.142)	(0.000)	(0.000)	(0.033)	(0.923)	(0.216)	(0.00)	
2	3	−0.468	7.720	1.309	−9.411	0.096	−1.121		0.62
		(0.205)	(0.000)	(0.000)	(0.042)	(0.721)	(0.033)	(0.57)	
2	4	−0.427	7.683	1.322	−7.623	0.106	−1.073		0.62
		(0.248)	(0.000)	(0.000)	(0.109)	(0.693)	(0.021)	(0.55)	
3	1	−0.569	7.540	1.364	−10.543	0.009	−0.050		0.59
		(0.123)	(0.000)	(0.000)	(0.026)	(0.972)	(0.629)	(0.00)	
3	2	−0.562	7.510	1.348	−10.825	0.020	−0.040		0.50
		(0.128)	(0.000)	(0.000)	(0.023)	(0.934)	(0.803)	(0.00)	
3	3	−0.513	7.661	1.340	−9.908	0.065	−0.559		0.61
		(0.163)	(0.000)	(0.000)	(0.031)	(0.806)	(0.047)	(0.07)	
3	4	−0.474	7.626	1.366	−7.895	0.068	−0.597		0.62
		(0.197)	(0.000)	(0.000)	(0.096)	(0.798)	(0.017)	(0.06)	
4	1	−0.545	7.601	1.363	−9.841	0.017	−0.210		0.59
		(0.140)	(0.000)	(0.000)	(0.039)	(0.948)	(0.182)	(0.00)	
4	2	−0.528	7.576	1.353	−9.874	0.034	−0.296		0.59
		(0.155)	(0.000)	(0.000)	(0.040)	(0.899)	(0.201)	(0.02)	
4	3	−0.445	7.747	1.291	−9.181	0.111	−1.382		0.62
		(0.230)	(0.000)	(0.000)	(0.047)	(0.680)	(0.026)	(0.70)	
4	4	−0.407	7.706	1.300	−7.580	0.122	−1.264		0.63
		(0.271)	(0.000)	(0.000)	(0.111)	(0.649)	(0.015)	(0.67)	

Note: The *p*-values for a two-tail test are presented below the coefficients; *p*(−1) is the *p*-value of a two-tail test that β_4 is different from −1.

The health trend (S) and discount rates (D) are as defined under table 4.

estimating equation (5) are shown in table 5. The results are reported for four separate measures of the discount and health care cost trend rate.

Under all measures, the regression coefficient of the estimated PRB liability divided by lagged book value (β_4) is, as expected, negative and is significantly different from zero at $p \leq 0.10$ in ten of the 16 cases reported. In nine of the 16 cases the coefficient is significantly different from −1 at a 10 percent level. The liability appears to be sensitive to the choice of parameters. When the discount rate is the firm's equity rate and the health care cost trend rate is assumed to be 4.5 percent, the β_4 coefficient is −1.775.

When the health care cost trend rate increase to 9 percent, the β_4 coefficient increases to -0.559.

The results suggest that the market forecasts of the health care cost trend rate does not exceed 7 percent because when the health care cost trend rate exceeds 7 percent, the β_4 coefficients are close to zero. The results also indicate that a discount rate based only on bond yields (pension rate and marginal borrowing rate), as suggested by the FASB (1985, par. 44; 1990, par. 31), is considered by the market as too low. Instead, the results appear consistent with market's assessment of a higher discount rate based on the return on equity or the cost of capital.

The results of estimating equation (6) are reported in table 6. The coefficient on PRB cash payments, β_1, is not significant (at a 5 percent level) in any of the 16 cases reported. The coefficient on the liability estimate, β_4, is again negative in all cases, and is statistically significant in eight of the 16 cases reported. β_4 is not statistically significant when the discount rate used is either the pension rate or the estimated marginal borrowing rate, which is consistent with the suggestion that the market uses a firm-specific discount rate (cost of equity or cost of capital). In addition, F-tests for the increase in R^2 values resulting from including the PRB liability were significant at the 5 percent level in ten of 16 cases and at the 10 percent level in 12 of the 16 cases. This indicates that the liability measure appears to be value-relevant to investors in addition to the cash information.

The results of estimating equation (7) for the second hypothesis are shown in table 7. The β_1 coefficients were not, as expected, significant at a 10 percent level for a two-tailed t-test in 14 of the cases reported.[17] The size of the coefficient, in general, lies between -4.182 and 3.203, which is significantly higher than the case when the liability measure is not included in the regression model ($\beta_{1,90}=-14.58$ in table 3).

The results reported in table 7 are somewhat puzzling because α_2, which was significant in tables 5 and 6, is no longer significant. Because this may be due to the high correlation between $y^1_{i,90}/y^1_{i,89}$ and the regression intercept, a collinearity analysis was conducted with the procedure of Belsley et al. (1980). The highest condition index was 17, which indicates medium collinearity at most.[18] Further analysis confirmed that most of the collinearity is from the correlation between the intercept term and $y^1_{i,90}/y^1_{i,89}$. The analysis was then repeated after excluding the book-over-lagged-book variable, $(y^1_{i,90}/y^1_{i,89})$, without any significant effect on the results and with the highest condition index reduced to 3.7.

VII. Summary

This study analyzes the valuation implications of firms' obligations to provide current and future retirees with postretirement benefits other than pensions. In particular, this study examined investors' assessment of disclosed PRB cash payments prior to SFAS No. 106, and whether a measure of the PRB obligation is regarded by investors as value-relevant in addition to the disclosed cash payments.

The main findings are that investors have changed their assessment of the PRB obligation over the period 1984–1990 from a position of underestimation during 1984–1986

---

[17] In the remaining two cases, β_1 is positive.

[18] According to Belsley et al. (1980), a condition index of below 10 indicates mild collinearity, and one of more than 30 indicates severe collinearity.

Table 6
Regressions to Ascertain the Value-Relevance of PRB Obligations
(1990 data, 231 observations)

$$\text{Model: } \frac{P_{i.90}}{y_{i.89}} = \alpha_0 + \alpha_1 \frac{x^o_{i.90}}{y_{i.89}} + \alpha_2 \frac{y_{i.90}}{y_{i.89}} + \beta_1 \frac{C_{i.90}}{y_{i.89}} + \beta_2 \frac{pex_{i.90}}{y_{i.89}} + \beta_3 \frac{upl_{i.90}}{y_{i.89}} + \beta_4 \frac{L_{i.90}}{y_{i.89}} + \eta^3_i.$$

S	D	α_0	α_1	α_2	β_1	β_2	β_3	β_4	$\bar{R}^2$
1	1	−0.391	7.695	1.211	−12.753	−9.253	0.179	−0.206	0.61
		(0.307)	(0.000)	(0.002)	(0.198)	(0.053)	(0.537)	(0.223)	
1	2	−0.402	7.705	1.211	−13.414	−9.340	0.173	−0.359	0.61
		(0.300)	(0.000)	(0.002)	(0.186)	(0.052)	(0.542)	(0.210)	
1	3	−0.405	7.787	1.259	1.683	−8.767	0.141	−1.582	0.62
		(0.283)	(0.000)	(0.001)	(0.853)	(0.063)	(0.609)	(0.050)	
1	4	−0.413	7.688	1.305	9.919	−7.565	0.107	−2.224	0.62
		(0.272)	(0.000)	(0.001)	(0.550)	(0.112)	(0.550)	(0.005)	
2	1	−0.377	7.695	1.192	−12.487	−9.322	0.191	−0.123	0.61
		(0.327)	(0.000)	(0.002)	(0.114)	(0.051)	(0.511)	(0.308)	
2	2	−0.386	7.711	1.181	−13.350	−9.523	0.187	−0.256	0.61
		(0.310)	(0.000)	(0.002)	(0.191)	(0.047)	(0.509)	(0.252)	
2	3	−0.422	7.761	1.274	−4.702	−8.841	0.135	−0.807	0.62
		(0.264)	(0.000)	(0.001)	(0.198)	(0.061)	(0.625)	(0.090)	
2	4	−0.439	7.665	1.334	1.545	−7.652	0.093	−1.170	0.62
		(0.244)	(0.000)	(0.000)	(0.867)	(0.109)	(0.736)	(0.040)	
3	1	−0.370	7.696	1.183	−12.070	−9.338	0.194	−0.068	0.61
		(0.337)	(0.000)	(0.002)	(0.089)	(0.051)	(0.501)	(0.288)	
3	2	−0.376	7.716	1.166	−12.711	−9.594	0.193	−0.153	0.61
		(0.323)	(0.000)	(0.003)	(0.067)	(0.045)	(0.496)	(0.212)	
3	3	−0.430	7.745	1.280	−6.471	−8.889	0.132	−0.428	0.62
		(0.256)	(0.000)	(0.001)	(0.343)	(0.060)	(0.632)	(0.081)	
3	4	−0.451	7.657	1.345	−2.057	−7.781	0.089	−0.622	0.62
		(0.232)	(0.000)	(0.001)	(0.792)	(0.103)	(0.749)	(0.050)	
4	1	−0.382	7.695	1.199	−12.683	−9.303	0.187	−0.158	0.61
		(0.321)	(0.000)	(0.002)	(0.138)	(0.051)	(0.519)	(0.310)	
4	2	−0.392	7.708	1.192	−13.547	−9.465	0.183	−0.311	0.61
		(0.302)	(0.000)	(0.002)	(0.119)	(0.048)	(0.520)	(0.275)	
4	3	−0.416	7.770	1.269	−3.579	−8.812	0.137	−1.077	0.62
		(0.271)	(0.000)	(0.001)	(0.659)	(0.062)	(0.620)	(0.045)	
4	4	−0.430	7.673	1.325	3.682	−7.600	0.097	−1.549	0.62
		(0.254)	(0.000)	(0.001)	(0.718)	(0.111)	(0.724)	(0.026)	

Note: The *p*-values for a two-tail test are presented below the coefficients.

to realization of the large size of the present value of the expected PRB liability during 1987–1990. When the firm-specific present value of the PRB obligation is estimated and the relation between market values and this estimate is analyzed, it is shown that the estimated PRB obligation is value-relevant to investors in addition to the cash payments, and that cash payments are not value-relevant when the accounting numbers are adjusted to reflect the estimated PRB obligation. This result is important because it shows that even a measure of the PRB obligation based on publicly available information is more informative than the cash information. Most firms are expected to release their private PRB information during fiscal year 1993. This is expected to be more informative than the PRB liability measure used here, and therefore more infor-

Table 7
Regressions to Ascertain the Superiority of Liability Information
(1990 data, 231 observations)

Model: $\frac{P_{i,90}}{y^{1}_{i,89}} = \alpha_0 + \alpha_1 \frac{x^{1}_{i,90}}{y^{1}_{i,89}} + \alpha_2 \frac{y^{1}_{i,90}}{y^{1}_{i,89}} + \beta_1 \frac{C_{i,90}}{y^{1}_{i,89}} + \beta_2 \frac{pex_{i,90}}{y^{1}_{i,89}} + \beta_3 \frac{upl_{i,90}}{y^{1}_{i,89}} + \eta^{4}_{i}.$

S	D	α_0	α_1	α_2	β_1	β_2	β_3	$\bar{R}^2$
1	1	0.724	7.765	−0.023	1.945	−5.200	−0.414	0.68
		(0.011)	(0.000)	(0.937)	(0.342)	(0.051)	(0.066)	
1	2	0.879	7.503	−0.092	−0.852	−4.784	−0.245	0.58
		(0.002)	(0.000)	(0.746)	(0.746)	(0.067)	(0.313)	
1	3	0.893	7.427	−0.045	−3.548	−4.625	−0.079	0.55
		(0.002)	(0.000)	(0.879)	(0.266)	(0.073)	(0.756)	
1	4	0.868	7.398	−0.040	−2.272	−4.627	−0.131	0.56
		(0.003)	(0.000)	(0.892)	(0.431)	(0.073)	(0.604)	
2	1	0.233	7.482	−0.050	3.203	−7.869	−0.080	0.70
		(0.327)	(0.000)	(0.788)	(0.070)	(0.007)	(0.578)	
2	2	0.902	7.691	−0.163	0.624	−5.150	−0.381	0.64
		(0.002)	(0.000)	(0.576)	(0.792)	(0.051)	(0.096)	
2	3	0.894	7.445	−0.050	−3.586	−4.603	−0.090	0.56
		(0.002)	(0.000)	(0.867)	(0.252)	(0.075)	(0.725)	
2	4	0.847	7.420	−0.033	−1.663	−4.671	−0.167	0.57
		(0.003)	(0.000)	(0.909)	(0.543)	(0.071)	(0.500)	
3	1	0.974	7.536	−0.322	2.712	−4.912	−0.755	0.74
		(0.001)	(0.000)	(0.246)	(0.229)	(0.076)	(0.001)	
3	2	1.005	8.098	−0.343	2.556	−6.191	−0.581	0.76
		(0.000)	(0.000)	(0.219)	(0.213)	(0.021)	(0.004)	
3	3	0.925	7.458	−0.076	−4.182	−4.388	−0.093	0.57
		(0.001)	(0.000)	(0.796)	(0.167)	(0.089)	(0.716)	
3	4	0.814	7.482	−0.019	−0.990	−4.761	−0.210	0.59
		(0.005)	(0.000)	(0.949)	(0.700)	(0.067)	(0.386)	
4	1	0.445	8.450	0.156	3.147	−6.259	−0.436	0.73
		(0.106)	(0.000)	(0.587)	(0.080)	(0.023)	(0.030)	
4	2	0.887	7.591	−0.124	−0.056	−4.947	−0.315	0.61
		(0.002)	(0.000)	(0.671)	(0.982)	(0.060)	(0.182)	
4	3	0.893	7.436	−0.047	−3.545	−4.619	−0.856	0.56
		(0.002)	(0.000)	(0.874)	(0.261)	(0.073)	(0.738)	
4	4	0.857	7.407	−0.037	−1.933	−4.648	−0.151	0.56
		(0.003)	(0.000)	(0.900)	(0.490)	(0.072)	(0.546)	

Note: The letter l above a variable in the model indicates adjustment for the PRB liability. The p-values for a two-tail test are given below the coefficients.

mative than the cash measure. These findings support the motivation behind the FASB's decision to adopt SFAS No. 106.

It appears that the PRB liability is highly sensitive to the choice of parameters—health care cost trend rate and discount rate—used to calculate it. The FASB tried to mitigate this problem by requiring some disclosure about the sensitivity of the PRB obligation and annual cost to the choice of the health care cost trend rate (SFAS No. 106, par. 74). A similar requirement for the discount rate used to calculate the obligation may be helpful to financial statement users. In addition, the results suggest that the market uses a measure of the health care cost trend rate close to the general inflation rate to estimate the obligation. It is also suggested that, unlike the discount rate suggested by the FASB in its Exposure Draft, a firm-specific measure of the return on equity or the cost of capital is used by investors to estimate the present value of the PRB obligation.

This research is limited by the fact that the estimate of the PRB liability cannot be fully validated. To date, no firm-specific demographic data are available to researchers, and only a small number of firms have adopted SFAS No. 106. When audited estimates are reported by firms in their financial statements, it will be possible to examine whether the actual reported liability is incrementally informative over the estimated liability used in this reserch. It would also be interesting to examine managers' choice of parameters to calculate the PRB obligation.

References

Amir, E., and J. Livnat. 1992. Recognition vs. disclosure: Does it make a difference? A study of other postretirement benefits. Working paper, Columbia University, New York.

Bacon, P., J. Kasper, G. De Lissovoy, S. Dicarlo, and J. Gabel. 1990. The employer response to the FASB proposal of accruing postretirement health benefits. *Benefits Quarterly* 6: 48–50.

Barth, M. E. 1991. Relative measurement errors among alternative pension asset and liability measures. *The Accounting Review* 66 (July): 433–63.

——, W. H. Beaver, and W. R. Landsman. 1992. The market valuation implication of net periodic pension cost components. *Journal of Accounting & Economics* 15 (March): 27–62.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics*. New York: McGraw-Hill.

Blue Book of Pension Funds. 1983, 1984–1985, 1985–1986, ERISA Benefits Funds, Inc., Washington, D.C.

Box, G. E. P., and G. M. Jenkins. 1976. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Espahbodi, H., E. Strock, and H. Tehranian. 1991. Impact on equity prices of pronouncements related to nonpension postretirement benefits. *Journal of Accounting & Economics* 14 (December): 323–46.

Financial Accounting Standards Board. 1984. *Statement of Financial Accounting Standard No. 81: Disclosure of Post Retirement Health Care and Life Insurance Benefits*. Stamford, CT: FASB.

——. 1985. *Statement of Financial Accounting Standard No. 87: Employers' Accounting for Pensions*. Stamford, CT: FASB.

——. 1990. *Statement of Financial Accounting Standard No. 106: Employers' Accounting for Post-Retirement Benefits other than Pensions*. Norwalk, CT: FASB.

Financial Executives Research Foundation. 1989. *Retiree Health Benefits: Field Test of the FASB Proposal*. Morristown, NJ: FERF.

Foster Higgins. 1988. *Third Annual Health Care Benefits Survey*. New York: A. Foster Higgins & Co. Inc.

Geisel, J. 1987. Court says firm can alter retiree health plan benefits. *Business Insurance* 21 (April): 6.

——. 1989. Burden of retiree health benefits growing. *Business Insurance* 23 (January): 2, 7.

General Accounting Office. 1989. Report to Congressional Requesters. *Employee Benefits*. Washington, D.C.: GAO.

Grant, J. 1991. Liabilities for non-pension retirement benefits—Valuation impact of expenditure disclosures. Working paper, Case Western Reserve University, Cleveland, OH.

Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47 (January): 153–61.

Ippolito, R. A. 1986. *Pensions, Economics and Public Policy*. Homewood, IL: Dow Jones-Irwin.

Landsman, W. 1986. An empirical investigation of pension fund property rights. *The Accounting Review* 61 (October): 662–91.

Lev, B. 1979. The impact of accounting regulation on the stock market: The case of oil and gas companies. *The Accounting Review* 54 (July): 485–503.

Mittelstaedt, F., and M. Wharshawsky. 1992. The impact of liabilities for retiree health benefits on share prices. Working paper, University of Notre Dame, Notre Dame, IN.

Ohlson, J. 1991. Earnings, book values, and dividends in security valuation. Working paper, Columbia University, New York.

U.S. Department of Commerce, Bureau of the Census. 1975. *Historical Statistics of the United States —Colonial Time to 1970*. Washington, D.C.: Government Printing Office.

———. 1986. *Money and Income of Households, Families and Persons in the United States*, Series P-60. Washington, D.C.: Government Printing Office.

———. 1989. *Statistical Abstract of the United States*, 109th ed. Washington, D.C.: Government Printing Office.

U.S. Department of Health and Human Services. 1988. *Vital Statistics of the United States 1986*. Hyattsville, MD: Government Printing Office.

U.S. Department of Labor. Pension and Welfare Benefits Administration. 1986. *Employer-Sponsored Retiree Health Insurance*. Washington, D.C.: Government Printing Office.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 725–747

Measuring Equilibrating Forces of Financial Ratios

Harry Zvi Davis
Baruch College, City University of New York
Yoram C. Peles
The Hebrew University of Jerusalem and
Hong Kong University of Science and Technology

SYNOPSIS AND INTRODUCTION: We test several categories of ratios—short-term liquidity, performance measures, earnings per share (EPS), capital structure, and the gross margin ratio—to determine if they have equilibrium values or follow a random walk. For ratios with an equilibrium value, the speed at which the ratio returns to equilibrium from out-of-equilibrium conditions is measured. Since equilibrating forces may differ with firm size, we also test for differences in adjustment speeds between small and large firms.

An accounting ratio may have an equilibrium value if management targets a certain ratio so that any deviation from the target causes management to initiate actions that will return the ratio to target. Also, although management may not be targeting a ratio, the interaction of management's actions with external market industry forces may lead to an equilibrium value. We investigate the total adjustment over time and then assess both the relative adjustment speed and the relative weights of the two main equilibrating forces: industry and management.

The statistical technique used allows each firm to have its own, unknown equilibrium value. In addition, we remove an important sampling bias in measuring the autocorrelation coefficient.

The results show that when firms experience a liquidity shock, equilibrating forces counterbalance a little more than a third of the shock in the next period. This finding suggests that firms' liquidity ratios have a fast adjustment to equilibrium values. EPS ratios also have a high adjustment rate to equilibrium value; about one-third to one-half of the shock is adjusted within one period. For performance measures (net operating income over sales or assets), for the equity to debt and gross margin ratios,

We thank S. Balsam, M. Benis, J. Bozewicz, B. Halperin, J. Kerstein, S. Lustgarten, M. Mellman, T. Ophir, the participants at the Baruch College seminar, and three anonymous reviewers for their helpful comments.

Submitted December 1990.
Accepted December 1992.

the findings imply a relatively long adjustment process to equilibrium values.

Supplementary tests suggest that smaller firms adjust their ratios to the optimal target more swiftly than large firms. Separating the ratios' total adjustment effect into industry and management components provides evidence that the adjustment rates differ for different industries. On average, the management adjustment is faster than the industry effect. A weighting measure suggests that both industry and management contribute a significant share to the total adjustment.

Tests of the predictive power of the model show that historical control of performance ratios is correlated with future actual performance. Also, firms that were acquired showed better than average control of their liquidity and performance ratios and showed a destabilization of their equilibrium EPS ratios. Finally, comparing the adjustment rate with a sample of firms from an earlier time period shows a stability of the adjustment behavior over time.

Information about the existence of equilibrium ratios and their adjustment speeds can help predict future values or events, and identify firms in special categories, for example, firms that will be acquired. It can also help in the evaluation of managerial actions insofar as managers have control over aspects of the financial ratios examined.

Key Words: *Financial ratios, Equilibrating forces, Equilibrium values.*

Data Availability: *Data used in this study may be obtained from the authors on request.*

THE article is organized as follows. We present the theoretical issues in section I. In section II we discuss various ratios and offer several explanations about the (non)existence of their equilibrium values. We describe the statistical technique and sample selection in section III and report the results in section IV. Supplementary hypotheses in section V test for a size effect; we also separate the total effect into industry and firm components, predict future performance, identify firms that will be acquired, and test for the stability of the adjustment rates obtained. The last section contains conclusions.

I. Theoretical Issues

The partial adjustment model of financial ratios postulates that firms have a constant coefficient of adjustment, β, by which their financial ratios adjust to target values. Lev (1969) tested the model by assuming that firms use industry averages as target values. Peles and Schneller (1989) relaxed Lev's assumption and allowed firm-specific, unobservable targets. Peles and Schneller make three important assumptions: (1) all firms have the same β, (2) the unobserved target is stable over time, and (3) there is no sampling bias in measuring the correlation coefficient.

This study improves upon Peles and Schneller's model by relaxing these assumptions. Different adjustment rates for different firms are allowed by dividing β into β^I, an industrywide adjustment rate, and β^M, a management rate. Because β^I differs across industries, firms in different industries may have different βs. Even within industries βs may differ because β^M is firm-specific. Also, rather than assume that the target ratio is stable, we model a system that allows changing targets. We control for the sampling size bias in measuring the autocorrelation coefficient then further extend the literature by using the adjustment rates to predict future events.

The specific research question tested is: Do equilibrating forces drive a ratio toward a target value, or does the ratio wander around in a Brownian motion? When an equilibrium ratio is found to exist, we measure the strength of the adjustment process and separate it into industry and management components. The ratios we study are the most commonly used ratios in accounting literature.

To begin the analysis, consider a ratio, V, whose value at time t is denoted V_t. In the absence of any economic shock (i.e., something that causes the ratio's value to change), the ratio remains at V_t. If, however, at $t+1$ the ratio's components are subject to a shock, the ratio shifts to a new value, denoted V_{t+1}. In the absence of any further economic shocks, V may either stay at V_{t+1} or revert to V_t. If V stays at V_{t+1}, the shock's effect is permanent and shifts the ratio toward a new level. This behavior is known as random walk. If, however, the ratio has a stable equilibrium value and the shock's effect is temporary, then V reverts to V_t. Reversion may occur in one period or in more than one period. β is the percentage of the remaining deviation adjusted each period.[1] That is, in the period following the shock, β of the shock is reversed; in the next period, another β of the remaining deviation from the equilibrium value is reversed, and so on.

Consider the correlation, ρ, between the first differences of the time series (ΔV_t, ΔV_{t+1}). If mean reversion occurs in period one, then $\rho = -\frac{1}{2}$; in a random walk, mean reversion does not occur, so , $=0$ (Beaver 1970, 67). More generally, the correlation of first differences may be expressed as in Peles and Schneller (1989, 529):

$$\rho(\Delta V_t, \Delta V_{t+1}) = -\frac{\beta}{2} \qquad 0 \leq \beta \leq 1, \tag{1}$$

where β measures the amount of mean reversion in each period following the shock.[2] At its maximum value, $\beta=1$, there is complete mean reversion; any shock in period one is followed by an equal and opposite equilibrating force in period two. At its minimum value, $\beta=0$, a shock in period one is not followed by equilibrating force; the ratio follows a random walk. Peles and Schneller (1989) show that, for $0<\beta<1$, reversion occurs over several periods. High (low) values of β (i.e., a high [low] negative serial correlation of ρ) imply strong (weak) equilibrating forces.

This analysis assumes that V has an equilibrium value. However, some ratios may not have equilibrium values. Therefore, one must first distinguish between those with and those without equilibrium values. Since each firm may have a unique equilibrium point, aggregation across firms may obscure the existence of equilibria. To address the problem, we use the statistical technique of Peles and Schneller (1989) to determine if a firm-specific equilibrium value (of V) exists; and if one does, the technique measures the speed at which V reverts to equilibrium from an out-of-equilibrium state.

---

[1] If the ratio is subject to a series of shocks over time, the net movement to equilibrium is β times the difference between the target and the actual ratio.

[2] Peles and Schneller (1989) use n, which represents the life used in a declining balance amortization. Since $\beta=1/n$, Lev's (1969) analysis of β is identical to Peles and Schneller's analysis.

Instability of the Target Ratio

Peles and Schneller (1989) derive equation (1) by assuming that a ratio's target value is stable. A more realistic model allows for shocks to both the ratio's target and actual values. In this section, we present such a model and discuss its empirical implications.

An accounting ratio may be affected by two types of changes: random shocks that affect only the observed value, and changes in the target value of the ratio. For example, assume that a cash ratio (e.g., cash to current assets) has a target level that is a negative function of the market interest rate. If interest rates fluctuate nonsystematically, the ratio's target fluctuates nonsystematically. In addition, the actual level of the ratio is subject to random shocks from unexpected inflows and outflows of cash generated by the operations of the firm. Thus, a cash ratio may incorporate both types of shocks.

Assume that the shocks to the actual value and to the target value are independent of each other and that both are zero-mean, independent, and identically distributed (i.i.d.) over time.[3] In any given period, equilibrating forces eliminate a fraction, γ, of the deviation between the target and actual ratios. It has been proved that,[4]

$$\rho(\Delta V_{t+1},\Delta V_t)=\frac{\gamma\sigma^2+\gamma(1-\gamma)\sigma^{*2}}{-2\sigma^2+\gamma\sigma^{*2}} \qquad 0\leq\gamma\leq 1, \tag{2}$$

where σ^2 and σ^{*2} represent the variance of the shocks to the actual value and the change in target value, respectively.

A finding of $\rho=0$ implies $\beta=\gamma=0$, a random walk model with no adjustments between the actual and the target ratio in response to changes in either the actual or the target ratio.[5]

Combining equations (1) and (2) gives,

$$\beta=\frac{2\gamma\sigma^2+2\gamma(1-\gamma)\sigma^{*2}}{2\sigma^2-\gamma\sigma^{*2}} \qquad 0\leq\beta\leq 1, 0\leq\gamma\leq 1. \tag{3}$$

If changes to the target are relatively insignificant ($\sigma^{*2}=0$), equation (3) simplifies to $B=\gamma$. In other words, the two adjustments are identical. If changes to the target are as large as changes to the actual ratio ($\sigma^{*2}=\sigma^2$), equation (3) simplifies to $B=2\gamma$. In other words, as changes to the target become more significant, β overstates γ, the actual total adjustment. When changes to the target become as large as changes to the actual ratio, the total adjustment is one-half β.

In summary, relaxing the assumption of a constant target has no effect on the finding of random walk. However, when there is a positive β, shocks to the target imply that β overstates the total adjustment.

---

[3] In this study, we assume that shocks are i.i.d. over time. The shocks can be partially dependent across firms. The observed dispersion of ρ reflects the diverse shocks and management responses by different firms. In section V, we measure the industry effect, that is, adjustments common to all firms in an industry. To the extent that this common factor exists, it lowers the dispersion of ρ and makes the results more reliable.

[4] See Davis and Peles (1992).

[5] The finding is also consistent with $\sigma^2=0$ and $\gamma=1$; there are no shocks to the actual value, yet all changes to the target are immediately adjusted for. This could explain a ratio over which management has immediate and total control (e.g., a depreciation ratio). However, it is inconsistent with any ratio whose actual value can be affected by external economic events.

Industry Versus Firm-Specific Effects

Two types of shocks may affect the actual ratios of a firm: (1) shocks affecting only a specific firm (e.g., a strike of the workers in the firm) and (2) shocks common to all firms in the industry (e.g., an increase in demand for industry products). Shocks may bring about (firm-specific) management actions (β^M) and/or equilibrating forces common to all firms in the industry (β^I).

The industry effect is a concurrent change common to all firms in an industry. It may reflect outside forces that are independent of firm actions, or simultaneous behavior by all firms in an industry. For instance, an industrywide decline in demand may result in inventory accumulations, and thus increase inventory ratios. The shock may be followed by management or industry equilibrating forces. A decline in profits that causes firms to leave the industry, which would lead to lower production and inventory levels in the remaining firms, is an example of an industry equilibrating force. A management-prompted decrease in production is an example of a management equilibrating force.[6] Different industries may have different β^I either because industry effects differ across industries or because managements in different industries equilibrate at different speeds.

Separating the total adjustment, β, into β^I and β^M is important when assigning responsibility for firm performance. Industrywide market forces operate on all firms in an industry and may be beyond management's control. However, management equilibrating forces are the result of management actions, and management should be accountable for these adjustments. Thus, when analyzing firm performance, β is relevant; when evaluating management performance, only β^M is relevant.

II. Hypotheses

Short-Term Liquidity

Short-term liquidity is an important concern of management. If the firm has insufficient current assets to satisfy current liabilities as they come due, the firm may be forced into bankruptcy. If the ratio of current assets to current liabilities decreases dramatically, managers must take measures to ensure that future cash flows are available to pay current liabilities as they mature. Even if the ratio does not shift, a decline in highly liquid assets may be disastrous, since usually neither receivables nor inventory can be used to pay current liabilities. Firms also need cash to pay bills that are not yet liabilities (e.g., payments for labor services). To the extent that these motives for holding liquid assets are important and vary from time to time, the ratios of liquid assets to short-term liabilities, and liquid assets to other current assets, may deviate from their targets.[7]

Other components of the current ratio are also subject to shocks. A sudden surge in sales may lead to a bulge in receivables that elevates the current ratio. Similarly, a surge in expected sales may lead the firm to increase its purchases, thus increasing current liabilities and probably inventory.

---

[6] If all firms in an industry identically reduce production, this force, even though it is internal, is an industry effect, since all the firms in the industry are responding to the shock identically. In this respect, any industrywide development is an industry effect.

[7] The firm is not trying to maximize the cash and liquid investments that it holds. Since the firm can usually earn a better return by investing in productive assets (if not, the firm could liquidate its excess assets, and the owners could directly invest in short-term marketable securities), excessive accumulations of liquid assets are not the best use of funds.

The equilibrating forces for short-term liquidity may be management- or industry-motivated. Since management targets certain ratio values below which the firm is believed to be in danger of being unable to satisfy creditors,[8] management will take actions to redress the balance in the ratios when a ratio falls below certain levels. For instance, management may delay capital investment to conserve cash or may issue long-term debt to raise cash. Alternatively, management may hasten the collection of receivables, but may find itself constrained by market forces. Similarly, reducing purchases may conserve cash, but may hurt the firm's sales.

Examples of industrywide forces that may affect the liquidity ratios of all firms in an industry or all firms in the market include a decline in demand that leaves firms with excess inventory. Or, if firms in an industry expect an increase (decrease) in demand, they may build up (reduce) inventory levels. Also, changes in interest rates affect the cost of holding inventory and other assets, which may cause a change in the target level of liquidity ratios. If the supply of a product used as an input changes (e.g., there is a good crop), firms that use the input may change their inventory levels. New competition may drive firms to increase credit to customers, or an economic downturn may cause customers to delay repayment.

In the accounting and finance literature, short-term liquidity and the components of current assets are measured by five ratios: current ratio, quick ratio, cash plus short-term investments to current assets, inventory to current assets, and current asset decomposition.[9] Since there are important reasons why managers may target liquidity ratios, we hypothesize that (1) liquidity ratios have equilibrium target values, and (2) shocks that affect these ratios raise forces that restore equilibrium values. The null hypothesis is that liquidity ratios have no stable target value; they follow a random walk.

Performance Measures

Management may attempt to maximize the firm's performance measures (such as profits or return on assets) either because the firm's goal is the maximization of owners' wealth or because management compensation is based on firm performance. However, actual and potential competitors of the firm may force these performance measures toward equilibrium values. A deviation from "normal" performance will cause the firm and competitors to react to restore the normality. For example, an increase in demand that improves a firm's performance will bring new investments into the industry, creating competition and lowering exceptional performance. Alternatively, in bad times, firms may leave the industry, thus improving the remaining firms' performance.

---

[8] See Foster (1986, 61).

[9] The current ratio is found in Horrigan (1965), Beaver (1966), Lev (1969), Deakin (1972), Edmister (1972), Elam (1975), Libby (1975), Altman et al. (1977), and Peles and Schneller (1989); the quick ratio in Horrigan (1965), Beaver (1966), Lev (1969), Deakin (1972), Edmister (1972), Chesser (1974), Elam (1975), and Peles and Schneller (1989); cash plus short-term investments to current assets figures in Beaver (1966) and Deakin (1972); inventory to current assets in Edmister (1972); and current asset decomposition in Lev (1968).

Lev's decomposition ratios are defined as $-\Sigma X_\alpha \log X_\alpha$, where X_α is a financial ratio. For the current asset decomposition, the summation is over all the items of current assets, and X_α equals the ratio of asset α over total current assets.

Management forces also may affect the firm's reported performance if management tries to smooth performance measures.

Two common performance measures in the literature are net operating income (NOI) to sales and NOI to assets.[10] We hypothesize that (1) performance ratios have equilibrium target values and (2) shocks that affect these ratios raise forces that restore equilibrium values.

Earnings Per Share

Of special interest are the performance measures of equity. The literature on the time-series properties of EPS ratios provides evidence that EPS does not follow a random walk.[11] Both Ball and Watts (1972) and Beaver (1970) find convincing evidence that EPS does not follow pure mean reversion. However, both studies find partial adjustment. Ball and Watts (1972, 671–72) find that the mean serial correlation coefficient for EPS is -0.198. This corresponds to an adjustment coefficient of 40 percent, that is, 40 percent of any shock is adjusted in the following year. Similarly, Beaver (1970, 86) concludes that "accounting rates of return also appear to be mean reverting, but the reversion takes place over several years." Freeman et al. (1982, 647–48) confirm Beaver's findings and conclude that "the observed accounting rates-of-return appear consistent with a mean-reverting process." Additional evidence may be found in Brooks and Buckmaster (1976), Collins and Hopwood (1980), and Brown et al. (1987).

Management has a high degree of control over EPS ratios.[12] Hence, we hypothesize that (1) EPS ratios have equilibrium target values, and (2) the equilibrating forces of EPS are strong, which implies a fast adjustment. Three types of EPS are examined: primary EPS excluding extraordinary items, primary EPS including extraordinany items, and fully diluted EPS excluding extraordinary items.[13]

Debt-Equity, Gross Margin, and Control Ratios

There are two opposing theories about the ratio of equity to debt.[14] Modigliani and Miller (1958) maintain that equity and debt are interchangeable; the other theory is that capital structure does matter.[15] Debt financing has the advantage of being tax deductible, but at the cost of increasing the risk (and cost) of bankruptcy. Therefore, the firm tries to maintain a ratio of equity to debt that optimizes the net benefit of tax savings minus the expected cost of bankruptcy.

[10] There are many performance measures in the literature. A recent measure is Tobin's Q ratio, which equals the firm's market value divided by the replacement cost of its assets (see, e.g., Morck et al. 1988). Although the technique in this study is also applicable to Tobin's Q, we concentrate on more traditional performance measures. NOI to sales is found in Horrigan (1965), Edmister (1972), and Elam (1975); NOI to assets is found in Horrigan (1965), Altman (1968), Lev (1969), Edmister (1972), Chesser (1974), Elam (1975), Libby (1975), Altman et al. (1977), and Peles and Schneller (1989).

[11] EPS ratios are different than dividend ratios in that earnings are subject to random shocks that can raise equilibrating forces, while dividends are set by management policy.

[12] For a detailed survey of income smoothing, see Ronen and Sadan (1981).

[13] Fully diluted EPS including extraordinary items is not examined because of insufficient data points.

[14] The ratio of debt to equity or equity to debt has been used in the literature extensively. See, for example, Horrigan (1965), Altman (1968), Lev (1969), Edmister (1972), Chesser (1974), Elam (1975), and Peles and Schneller (1989).

[15] These opposing theories are discussed in Miller (1977). Miller claims that the Modigliani-Miller (1958) theory is valid even in a world with corporate taxes. "I will argue that even in a world in which interest payments are fully deductible in computing corporate income taxes, the value of the firm, in equilibrium will still be independent of its capital structure" (1977, 262).

If firms' capital structures behave according to the Modigliani-Miller (1958) theory, random shocks are not adjusted; therefore, we predict that the debt-equity ratio follows a random walk ($\beta=0$). If firms do not behave according to the Modigliani-Miller theory, then we predict that (1) debt-equity ratios have equilibrium target values, and (2) shocks that affect these ratios raise forces that restore equilibrium values.

Many studies in the accounting and economics literature claim that the gross margin is stable over time.[16] Thus, we predict that (1) the gross margin ratio has an equilibrium target value, and (2) shocks that affect this ratio raise forces that restore equilibrium value.

We do not expect all ratios to have equilibrium values. If one part of a ratio is measured in current dollars (e.g., sales) and the other item is measured in historical dollars (e.g., fixed assets), the ratio may be affected by a trend. Similarly, if the two parts of the ratio have no relationship to each other, there may be no equilibrating forces affecting the ratio. We choose four ratios from the accounting and finance literature—equity to fixed assets, sales to fixed assets, sales to equity, and retained earnings to total assets[17]—for which there is no convincing rationale that they have an equilibrium value. Thus, we hypothesize that the control ratios show no equilibrium adjustments.

Formally, for each ratio we hypothesize:

H1: $\beta \leq 0$ versus $\beta > 0$.

Moreover, we quantify β for each ratio.

III. Methodology

Measuring Adjustment Rates

The Peles-Schneller (1989) technique determines which value of β best describes the data by a multistep process. First, an arbitrary value of β is chosen. For this value, the theoretical ρ is calculated as $-\beta/2$. Second, for each firm, i, the actual $\rho=r_i$ is calculated. After adjusting both ρ and r_i for statistical reasons (see below), ρ is subtracted from r_i, which yields a z score. If the chosen β is correct, the z distribution will be standard normal. Third, a χ^2 test is used to compare the z distribution to a standard normal distribution. The higher the χ^2, the greater the probability level that the β is *not* the correct empirical value. Fourth, this procedure is repeated for different β values to produce a χ^2 for each β. In effect, for each β, we test the hypothesis that the z distribution is $N(0,1)$. Fifth, the β producing the minimum χ^2 value is chosen as the point estimate of β, since the minimum χ^2 implies that the β at that level produces a distribution closest to $N(0,1)$.[18]

The formulae used by Peles and Schneller (1989) follow. After r_i is calculated for each firm (by using a time series of length m), the r distribution is standardized. Since r_i is calculated from a limited number of periods and since $E(r_i)\neq 0$, Hotelling's (1953) significance test for small samples (which incorporates Fisher's correction for $E[r_i]\neq 0$) is used to normalize r_i. Accordingly, z_i scores are calculated as follows:

---

[16] See Hall and Hitch (1939), Saxton (1941), Kaplan et al. (1958), Cyert and March (1963), Eckstein and Fromm (1968), Silberston (1970), Eckstein and Wyss (1972), Wood (1975, 61), and Wilder et al. (1977).

[17] The first three ratios are found in Horrigan (1965) and Elam (1975); the third ratio is also found in Beaver (1966); the last ratio is found in Altman (1968) and Altman et al. (1977).

[18] This is the minimum χ^2 procedure of Mood et al. (1974, 286).

$$z_i=(Y_i^*-v_i^*)\sqrt{(k-1)}, \tag{4}$$

where:

$$Y_i^*=Y_i-\frac{3Y_i+r_i}{4(k-1)}, \tag{5}$$

and

$$Y_i=0.5\ln\left(\frac{1+r_i}{1-r_i}\right), \tag{6}$$

where:

r_i = the observed serial correlation of the *i*th firm,

k = the length of the time series, which equals the number of $(\Delta V_t, \Delta V_{t+1})$ pairs (i.e., $k-1=m-3$ degrees of freedom[19]), and

v^* is defined in a similar fashion as Y^* in terms of the hypothesized correlation coefficient.

The χ^2 goodness-of-fit test is used to test for normality of the z distribution by first dividing the standard normal distribution space into $n=18$ intervals,[20]

$$(-\infty,-2.00],(-2.00,-1.75],\ldots,(1.75,2.00],(2.00,+\infty). \tag{7}$$

Within each interval, the observed frequency distribution of the z_i scores for each set, f_n, is compared to the standard-normal frequency, F_n, by constructing the statistic,

$$\chi^2=\sum_{n=1}^{18}\frac{(f_n-F_n)^2}{F_n}. \tag{8}$$

Adjustment for the Sampling Bias

Equations (4) through (8) assume that there is no sampling bias. However, Kendall et al. (1983, 550–54) point out the existence of an autocorrelation sampling bias. They show (eq. 48.26, p. 552) that the correct formula for the expected serial correlation, $E(r_i)$, of a sample of k terms is,

$$E(r_i)=\rho+\frac{(1+\rho)(4\rho^2-2\rho-1)}{k-1}. \tag{9}$$

For a large time-series sample size, the equation simplifies to:

$$\underset{k\to\infty}{E(r_i)=\rho}, \tag{10}$$

which is Peles and Schneller's (1989) assumption. However, for small sample sizes—Peles and Schneller and this study have only 18 pairs of time-series data—the bias cannot be ignored. To correct for the bias, we calculate a transform of the hypothesized correlation coefficient as follows:

---

[19] There are m years of data, V_t, which generate $m-1$ annual differences, ΔV_t, which generate $m-2$ pairs of differences $(\Delta V_t, \Delta V_{t+1})$.

[20] There are 18 intervals of size 0.25 standard errors in the relevant space.

$$v(adj)_i^* = v_i^* + \frac{(1+v_i^*)(4v_i^{*2} - 2v_i^* - 1)}{k-1}. \qquad (11)$$

Equation (4) then becomes:

$$z_i = [Y_i^* - v(adj)_i^*]\sqrt{(k-1)}\,. \qquad (12)$$

This modification ensures that both the observed and the hypothesized coefficients have the same bias.[21]

Evaluation of Methodology

The modified Peles-Schneller (1989) methodology offers four benefits over other tests used in the literature. One, the method discriminates not only between two opposing models[22] (mean reversion vs. random walk), but also measures the rate at which deviations are reversed. Two, it is possible (see section V) to separate the industry adjustment from the firm-specific adjustment.[23] Three, because it uses the significance of the total distribution, the technique provides a more powerful test than time-series studies that test hypotheses based on the significance of a single data point (see, e.g., Ball and Watts 1972). However, since the test compares two distributions, actual and theoretical, it requires a lot of data. Four, the procedure does not require that all firms in an industry have the same equilibrium value for a ratio. In contrast to the restrictions in Lev (1969), each firm may have its own, unknown equilibrium value.

Sample Selection

The appendix defines each ratio. The data are from the 1972–1991 COMPUSTAT annual tapes. For each ratio, any firm that does not have 20 years of data is eliminated. Also, any four-digit SIC industry with less than two firms is dropped. For the industry tests, any industry with less than three firms is excluded. For the testing of small versus large firms, no industries are excluded. About 1,000 firms have the required information. Data for the supplementary tests are also taken from the COMPUSTAT 1972–1991 research tapes and the 1963–1982 annual tapes.

IV. Results

We begin by analyzing the results of two ratios in depth. Next, we summarize the findings of the other ratios. For each ratio, the χ_q^2 statistic is calculated for all $\beta_q, q=(0,\ldots,.50)$ in 0.025 increments. The minimum χ_q^2 is then determined. Figure 1 shows the χ_q^2 significance levels for $\beta_q \in (0, 0.5)$ for the quick ratio. For 20 years of data (17 d.f.), $\chi_q^2 \leq 27.6$ implies that the null hypothesis cannot be rejected at the 5 percent sig-

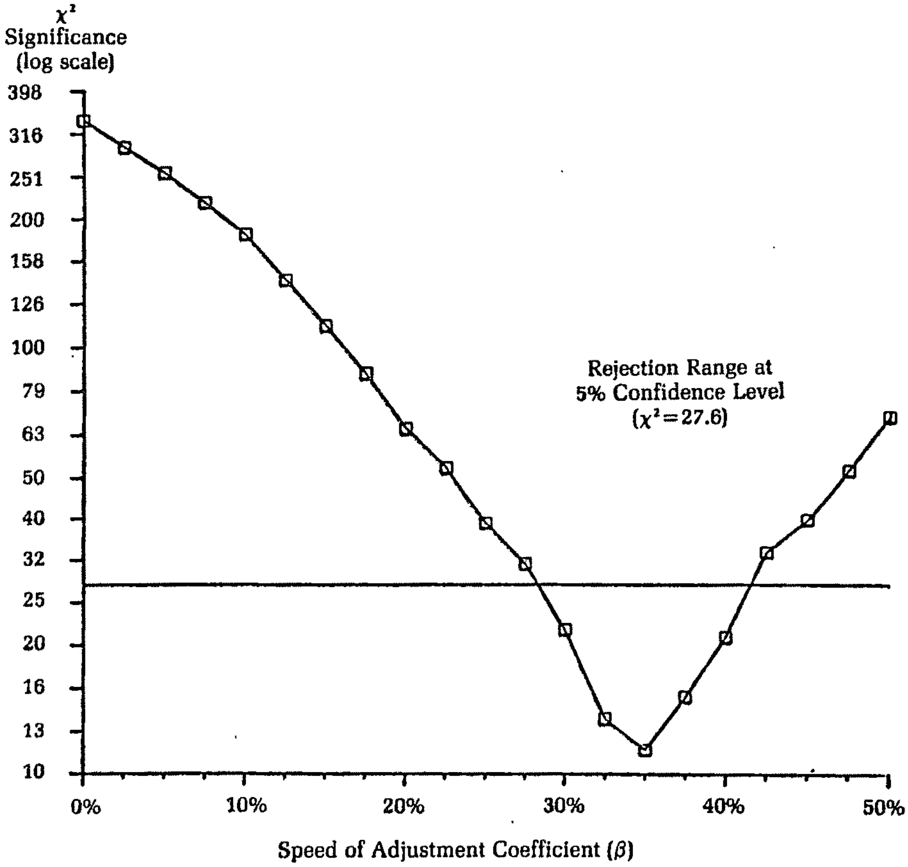
---

[21] We can correct for the bias only when using the Peles and Schneller (1989) technique to calculate the chi-square for each β_q. When we compare two populations (i.e., large vs. small firms), the identical bias is present in both populations and, thus, does not affect the difference tests. The point estimate for β we show in each population is the unbiased estimate of β calculated for each population separately with the modified Peles-Schneller technique.

[22] Beaver (1970).

[23] This is impossible with the approach of Lev (1969).

Figure 1
Quick Ratio

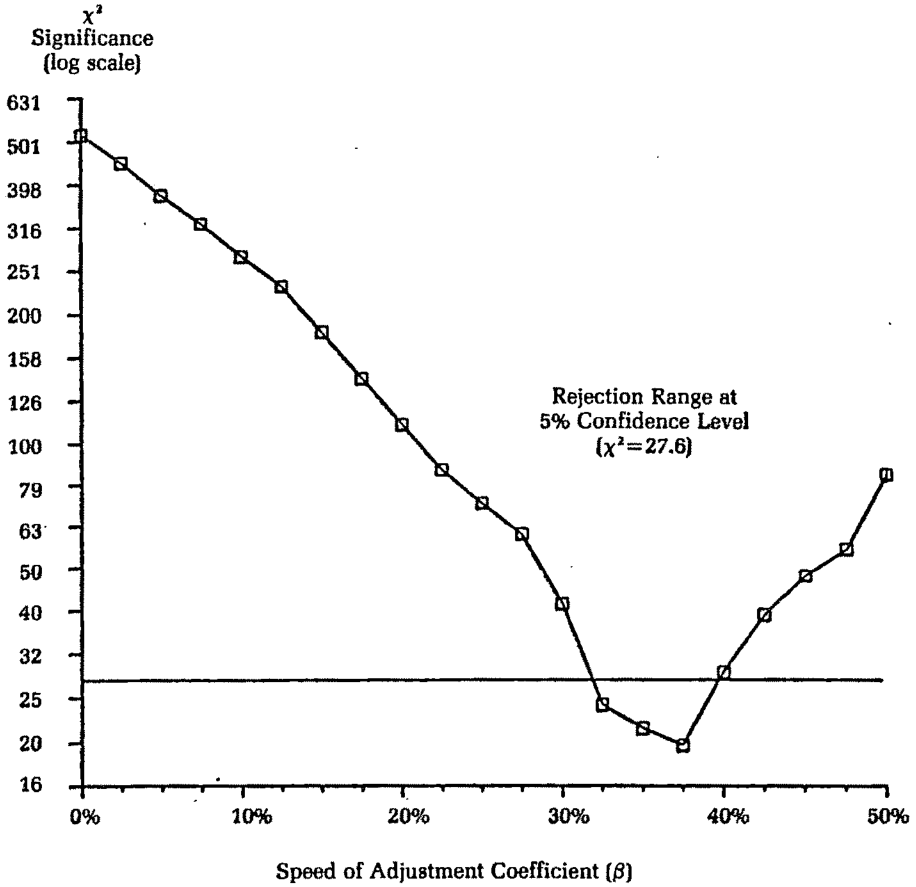


nificance level. The higher (lower) the significance level, the less (more) likely that the β_q is correct. The lowest significance level represents the point estimate of β_q. β_q coefficients for q between 30 and 40 percent are significant at the 5 percent level. The results are robust; changing to the 1 percent significance level ($\chi^2_q=24.8$) does not change the range of significant β_qs. For q less than 30 or more than 40 percent, there is a sharp increase in significance levels. The χ^2_q at $\beta_q=0$ percent, which represents a random walk model, is 340.

Two conclusions can be reached. First, the range of β_q is 30 to 40 percent with a point estimate of 35 percent. The smoothness of the curve is also reassuring; it implies that the change in significance values is continuous over the total range of β_q. Thus, testing β_q values in increments of less than 0.025 would smooth the curve, but would not alter the results. Second, since positive β_q values are statistically significant, ρ must be negative, which allows rejection of the null form of hypothesis H1 with the conclusion that the ratio has a stable target.

Figure 2 shows the χ^2_q significance levels for β_q of primary EPS excluding extraordinary items. Adjustment coefficients between 32.5 and 40 percent are significant at the 5

Figure 2
Primary Earnings per Share
(excluding extraordinary items)



percent level. The results are robust; they do not depend on the significance level chosen. Switching to a 10 percent significance level or even to a 0.1 percent significance level, barely changes the range. For all other β_q ($\beta_q \leq 30$ percent or $\beta_q \geq 42.5$ percent) there is a sharp increase in significance levels. The χ_q^2 for the random walk model is more then 520. Again, two conclusions can be reached: (1) the range of β_q is 32.5 to 40 percent, with a point estimate of 37.5 percent, and (2) the ratio has a stable equilibrium.

Adjustment Coefficients of Ratios

Table 1 shows a dichotomy between the liquidity and EPS ratios, whose β is 35 to 50 percent, and the other ratios, whose β is 2.5 to 7.5 percent. The five liquidity ratios provide very similar results. The point estimates of β vary between 35 and 50 percent. All χ_q^2 significance levels are low, which implies that the value cannot be rejected at reasonable significance levels. The χ_q^2 significance levels of random walk, 319 through 538, show that null form of hypothesis H1 is rejected at all reasonable levels.

Table 1

Point Estimates of Adjustment Coefficients, 1972–1991

	Number of Firms	*Point Estimate*		*Random Walk (β=0%)*
		Adjustment Coefficient β	*Chi-Square*	*Chi-Square*
Liquidity Measures				
Current Ratio	1,092	37.5%	24.8	508.3
Quick Ratio	941	35.0%	11.4	340.5
Cash Plus Short-Term Investments/Current Assets	943	37.5%	11.0	389.3
Inventory/Current Assets	875	37.5%	18.3	318.9
Current Asset Decomposition	850	50.0%	16.4	538.2
Performance Measures				
Net Operating Income/Sales	1,171	7.5%	15.5	21.6
Net Operating Income/Assets	1,171	2.5%	16.2	26.5
Earnings Per Share				
Primary - Excluding Extraordinary Items	1,242	37.5%	19.7	520.8
Primary - Including Extraordinary Items	1,245	47.5%	13.1	939.2
Fully Diluted - Excluding Extraordinary Items	998	35.0%	14.4	386.2
Capital Structure				
Equity/Debt	862	5.0%	13.0	22.0
Gross Margin Ratio	1,052	2.5%	26.0	31.4

Note: β measures the percentage of mean reversion of a ratio. The point estimate represents the minimum χ^2_q for all possible q values between zero and one at intervals of 2.5 percent. χ^2_q=27.6 at the 5 percent significance level with 17 degrees of freedom. The appendix provides detailed descriptions of the ratios.

The point estimates for the performance measures are 2.5 to 7.5 percent, which also rejects the null form of hypothesis H1. For the NOI to assets ratio, the random walk model can almost be rejected at the 5 percent level and definitely can be rejected at the 1 percent significance level (χ^2=24.8). This is weak evidence that these ratios do not follow a random walk.

The point estimates for the three variants of EPS are β=35 to 47.5 percent. Since the χ^2_q significance levels of a random walk are all more than 386, the data imply that EPS, however it is defined, does not follow a random walk behavior. The fastest EPS adjustment (and the highest χ^2_q for the random walk model) is for earnings including extraordinary items (47.5 percent compared to 35 percent and 37.5 percent). This is intuitively appealing because extraordinary items are vulnerable to manipulations and can be used by management to adjust income. The fully diluted EPS has the lowest adjustment speed (35 percent compared to 37.5 percent and 47.5 percent) and the lowest χ^2_q for the random walk of the EPS ratios. This is because fully diluted EPS is subject to factors beyond management's control. For instance, the extent of dilution in the treasury stock method depends on the current market price of the firm's shares, a variable that management cannot control.

It is interesting to compare the EPS results with the NOI measures and the gross margin ratio. EPS, however defined, has a very strong annual adjustment factor; between one-third and almost one-half of any deviation is adjusted annually, while for the other three measures, only 2.5 to 7.5 percent of the deviation is adjusted annually. The difference can be explained by the importance the economic community and investors assign to EPS. This importance is reflected in the financial press, which prominently publishes each firm's price/earnings ratio, a ratio based on EPS.

For the equity-to-debt ratio, the point estimate is $\beta=5$ percent, which rejects the null form of hypothesis H1. This result is not consistent with the Modigliani-Miller (1958) theory that firms are indifferent to their capital structure; over a broad range, shocks that change a firm's actual capital structure are counteracted by management, albeit over a long period. However, the random walk model cannot be rejected.

The point estimates for the gross margin ratio is 2.5 percent, which rejects the null form of hypothesis H1. Furthermore, the random walk model can be rejected at the 5 percent level.

The finding that β is non-negative ($\rho\leq 0$) for all ratios is consistent with shocks to the actual firm ratio. In no case is β negative ($\rho>0$), a finding that would be consistent with shocks to the target ratio. This suggests that, for the ratios studied, the important changes are caused by random shocks to the actual ratio; target changes are quantitatively less important.

Finally, the four control ratios—equity to fixed assets, sales to fixed assets, sales to equity, and retained earnings to total assets—are not reported in the table since they each show a maximum likelihood β_q of 0 percent, which implies a random walk process. In all four cases, the random walk model, although not statistically significant ($\chi^2_q>70$), is a better description of the behavior over time than mean reversion. Hence, for these ratios, we cannot reject the null form of hypothesis H1 that $\beta\leq 0$.

V. Supplementary Hypotheses

Size Effect

If a ratio has an equilibrium value, the strength of the combined industry and management equilibrating forces may differ across firms. Peles and Schneller (1989) suggest that size may explain this difference. On the one hand, large firms have more resources and better access to capital markets than smaller firms. Their size may allow them to equilibrate faster than small firms. On the other hand, small firms may have more flexibility and less bureaucracy than large firms. Small firms may also suffer a larger penalty for being out of equilibrium than large firms. If so, they have a greater incentive to adjust to equilibrium than large firms. However, since firms can adjust EPS quite easily, we do not expect a difference in EPS ratios between large and small firms. We test these competing hypotheses by comparing the β coefficients of small and large firms. Formally, for each ratio, we hypothesize:

H2: The β of small firms = the β of large firms, versus β of small firms $\neq\beta$ of large firms.

We use a *t*-test to compare the βs for small versus large firms. Large firms are defined as those with above-average sales; small firms are defined as those with below-average sales.[24]

---

[24] The results are similar when splitting the firms according to either sales or assets. Mann-Whitney *U*-test results are similar to the parametric *t*-test results and are not reported.

Table 2

Test of Mean Adjustment Coefficients for Small Versus Large Firms

	Number of Firms	*Point Estimate of β*	*t-Value for Difference*	*Significance Level*
Current Ratio				
Small Firms	874	37.5%	3.89	0.000
Large Firms	218	25.0%		
Quick Ratio				
Small Firms	758	38.0%	3.13	0.002
Large Firms	183	22.0%		
Cash Plus Short-Term Investments/Current Assets				
Small Firms	760	37.5%	1.94	0.053
Large Firms	183	25.0%		
Inventory/Current Assets				
Small Firms	702	42.5%	3.90	0.000
Large Firms	173	20.0%		
Current Asset Decomposition				
Small Firms	681	50.0%	3.00	0.003
Large Firms	169	45.0%		
Net Operating Income/Sales				
Small Firms	941	12.5%	3.57	0.000
Large Firms	230	2.5%		
Net Operating Income/Assets				
Small Firms	941	2.5%	not significant	not significant
Large Firms	230	2.5%		
Primary EPS - Excluding Extraordinary Items				
Small Firms	996	37.5%	not significant	not significant
Large Firms	246	37.5%		
Primary EPS - Including Extraordinary Items				
Small Firms	1,000	47.5%	not significant	not significant
Large Firms	245	47.5%		
Fully Diluted EPS - Excluding Extraordinary Items				
Small Firms	785	35.0%	not significant	not significant
Large Firms	213	37.5%		
Equity to Debt				
Small Firms	675	5.0%	2.63	0.009
Large Firms	187	0.0%		
Gross Margin Ratio				
Small Firms	849	2.5%	6.21	0.000
Large Firms	203	0.0%		

Note: β measures the percentage of mean reversion of a ratio. The appendix provides detailed descriptions of the ratios.

Table 2 shows the point estimate of β for each size partition and ratio. One can observe two patterns. All the variants of EPS and NOI to assets show the same adjustment for small and large firms. This indicates that the ability to adjust for shocks in EPS does not depend on the size of the firm. For all other ratios, small firms show higher β coefficients than large firms. For liquidity ratios and NOI to sales, small firms' β is about 10 percent larger than that of large firms. For the equity-to-debt and gross margin ratios the differences are 2.5 to 5 percent. The *t*-tests to compare the βs of large and small firms are significant for seven of the ratios at the 0.01 level and for eight ratios at the 0.05 level.[25] For these ratios, we reject the null form of hypothesis H2 that the β of small and large firms is the same; small firms have a faster adjustment than large firms.

Testing for Industry Versus Management Effects

We separate the total adjustment into industry and management forces. The industry effect incorporates all forces common to firms in an industry, including activities taken in concert by managements in response to industrywide shocks. The speed at which industry forces operate is a function of the technology of the investment and disinvestment of each industry; this technology differs for different industries.[26]

A direct test of the effect of industry versus management separates β into β^I and β^M components and compares β to β^M. To derive β^M (equal to $\beta-\beta^I$), we estimate β^I by calculating a series of average first differences across firms for each four-digit SIC code industry in each year *t*, where $t=1,19$. The actual formula for the average first difference of each industry, *IND*, in year $t, \Delta V_t^{IND}$, is:

$$\Delta V_t^{IND}=\sum_{i=1}^{n}\frac{\Delta V_t^i}{n}, \tag{13}$$

where ΔV_t^i is the first difference of the *i*th firm in year *t*, and there are *n* firms in the industry. This step produces 19 average first differences in each industry. For each firm, the net of industry first difference $[\Delta V_t^i(net)]$ is then calculated by subtracting the average industry first difference from the firm first difference,

$$\Delta V_t^i(net)=\Delta V_t^i-\Delta V_t^{IND}. \tag{14}$$

The $\Delta V_t^i(net)$ terms are then tested, with the modified Peles and Schneller (1989) technique to calculate β^M. Industry averages are tested separately to calculate β^I. By comparing β^M to the total β, which combines both industry and management adjustment, we assess the relative adjustment speeds of management versus the total adjustment. For statistical significance testing, a matched *t*-test compares the total β distribution to β^M.

Since β is affected by both management manipulation and external industry forces, it is a weighted average of β^I and β^M.

$$\beta=\omega\beta^I+(1-\omega)\beta^M, \tag{15}$$

so that,

$$\omega=\frac{\beta-\beta^M}{\beta^I-\beta^M}, \tag{16}$$

---

[25] In a recent article, Fama and French (1988) found that stock prices are mean reverting. The mean reversion is stronger for small firms than for larger ones, which is in line with the findings here that small firms have a higher β than large firms.

[26] To test whether the ratio of capital to sales can serve as a proxy for technology, we tested capital-intensive versus noncapital-intensive companies, defining capital-intensive companies as those having above-average capital to sales. We observed no significant differences in estimates of rate of adjustment.

where ω and $(1-\omega)$ are, respectively, the contributions of β^I and β^M to the total adjustment.

We hypothesize that management equilibrating forces are more powerful than industry forces. Formally, we propose:

H3: $\beta^M \leq \beta$, versus $\beta^M > \beta$.

Table 3 presents the point estimates of β, β^M, β^I, ω, and matched *t*-test results. The low χ^2 scores suggest that all point estimates are a good fit for the data. For all the ratios, $\beta^M > \beta^I$, which suggests that management adjustment is faster than the industry effect. The ω values show that both industry and management forces contribute to the equilibrating effect. For some ratios, the industry effect is stronger; for others, the management effect is stronger. The high ω for EPS ratios can be explained by management behavior in response to shocks that affect all firms in an industry. However, it should be noted that the calculation of ω is very sensitive to small changes in the point estimates. Hence, ω should be interpreted more as a general indication of magnitude rather than as a precise estimate. A matched *t*-test of the total β distribution with the β^M distribution rejects hypothesis H3 for all ratios except the quick ratio.

Prediction

Two tests use historical β ratios to predict future events. These tests also provide external validation for the measurement of β. The first test uses the results obtained to predict future performance of firms. Firms that show better control of their performance ratios by adjusting deviations faster should have higher values of these ratios in the future. Therefore, the test is whether the autocorrelation of first differences of the NOI ratios for the 19 years (1971–1990), ρ^{19}, is correlated with the ratio value in the twentieth year, V^{20}. Since β implies a negative ρ, the sign of the coefficient of correlation should be negative. Formally, we propose:

H4: $Corr(\rho^{19}, V^{20}) \geq 0$ versus $Corr(\rho^{19}, V^{20}) < 0$, for both the NOI to sales and NOI to assets ratios, where *Corr* is the correlation coefficient.

The results obtained are consistent with the hypothesis. The coefficient of correlation is -0.091 for NOI to sales and -0.095 for NOI to assets. In both cases, the results are significant at the 0.001 level (1,088 firms).[27] We thus reject the null hypothesis and conclude that, for NOI ratios, historical data can be used in predicting future performance.

A second test for the predictive ability of the adjustment rate is to examine the β of firms that we expect to have different speeds of adjustment than control firms. Firms that were acquired or merged were compared to control firms for short-term liquidity, performance, and EPS ratios.[28] Formally, we propose for both short-term liquidity and performance ratios:

H5: The β of merged firms $\leq$ the β of control firms, versus β of merged firms $>$ β of control firms.

---

[27] The nonparametric Kendall and Spearman correlation coefficients are also negative, with significance levels of 0.001 and 0.002, respectively.

[28] We also compared control firms to firms that eventually went bankrupt or liquidated. The data and analysis are the same as for the acquired firms. However, because of the long period required by our statistical technique, we did not have a large enough sample size. Still, the results obtained are in the right direction; distressed firms have less stability in their liquidity and performance ratios than control firms. Representative results of the distressed versus control firm ratios are: current ratio 35 percent versus 40 percent; quick ratio 27.5 percent versus 42.5 percent; NOI to assets 0.0 percent to 17.5 percent; and, primary EPS excluding extraordinary items 42.5 percent versus 45 percent.

Table 3
Point Estimates of Total, Firm-Specific, and Industry Adjustment Coefficients, and Statistical Tests

	Current Ratio	*Quick Ratio*	*Cash Plus Short-Term Investments/ Current Assets*	*Inventory/ Current Assets*	*Current Asset Decomposition*	*Net Operating Income/ Sales*	*Net Operating Income/ Assets*	*Primary EPS Excluding Extraordinary Items*	*Primary EPS Including Extraordinary Items*	*Fully Diluted EPS Excluding Extraordinary Items*	*Gross Margin Ratio*	*Equity/ Debt*
Point Estimate of Total Adjustment Coefficient, β	37.5%	37.5%	40.0%	32.5%	50.0%	2.5%	2.5%	35.0%	50.0%	35.0%	2.5%	5.0%
(Chi-Square)	(24.6)	(14.4)	(15.5)	(18.9)	(13.8)	(11.8)	(16.3)	(13.5)	(14.2)	(9.5)	(28.5)	(16.0)
Number of Firms	866	714	714	677	650	941	941	1,008	1,012	795	823	686
Point Estimate of Net of Industry Firm-Specific Adjustment Coefficient, β^M	45.0%	37.5%	45.0%	45.0%	57.5%	17.5%	22.5%	52.5%	72.5%	52.5%	7.5%	15.0%
(Chi-Square)	(10.0)	(17.5)	(17.1)	(15.5)	(18.4)	(12.6)	(19.4)	(15.0)	(79.7)	(20.8)	(22.3)	(17.8)
Point Estimate of Industry Adjustment Coefficient, β^I	27.5%	27.5%	20.0%	25.0%	45.0%	0.0%	0.0%	30.0%	40.0%	22.5%	0.0%	0.0%
(Chi-Square)	(16.3)	(7.5)	(10.9)	(13.1)	(20.0)	(6.1)	(10.5)	(17.9)	(17.8)	(13.8)	(17.3)	(25.8)
Number of Industries	194	189	187	179	174	208	208	212	213	178	209	155
Weight of Industry Effect, ω	43%	25%	20%	62%	60%	85%	89%	78%	70%	58%	67%	67%
t-Value for Difference Between β and β^M	2.08	0.48	3.27	5.12	3.92	7.61	8.24	8.15	11.41	8.16	5.86	4.65
Significance Level	0.019	0.630	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Degrees of Freedom	865	713	713	676	649	940	940	1,007	1,011	794	822	685

Note: β measures the percentage of mean reversion of a ratio. The point estimate represents the minimum χ^2 for all possible β, β^M, and β^I values between zero and one at intervals of 2.5 percent. $\chi^2=27.6$ at the 5 percent significance level with 17 degrees of freedom. The appendix provides detailed descriptions of the ratios.

Of special interest is the behavior of EPS ratios. On the one hand, investors assign high importance to EPS ratios. On the other hand, management can easily manipulate these ratios. Before ownership changes, management prefers to show a rising EPS. Any such manipulation works to destabilize EPS. Hence, for EPS ratios, we propose:

H6: The β of merged firms $\geq$ the β of control firms, versus β of merged firms $<$ β of control firms.

The control firms are taken from all firms with 20 years of data on the COMPUSTAT tapes, 1972–1991. The sample of merged and acquired firms are from the COMPUSTAT research tapes for the same years (code 1). β_i is calculated for each firm from 12 years of data, 1980–1991 for control firms and the last 12 years prior to deletion for merged firms.[29] For both hypotheses H5 and H6, we use a Kolmogorov-Smirnov two-sample test to compare the distributions of the control firms with the merged firms.[30]

Table 4 presents the results. Panel A shows the results for hypothesis H5 (liquidity and performance ratios), and panel B shows the results for hypothesis H6 (EPS ratios). In panel A, firms that were acquired or merged show higher βs than the control firms for all ratios except the quick ratio, which is identical for both groups. For two of the liquidity ratios and one of the performance ratios, the null hypothesis H5 can be rejected at the 0.05 significance level. Thus, we find some evidence that these firms were managed better than the average firm. In panel B, firms that were acquired or merged show lower βs than the control firms for all EPS ratios. Furthermore, the differences are all significant at the 0.00 level. This provides evidence that the equilibrium EPS of merged firms is less stable than that of control firms.

Stability of the Adjustment Rate

The usefulness of a parameter depends on its stability over time. Testing for the stability of the adjustment rate was done by replicating the test of hypothesis H1 for firms over the time 1963–1982. Half of this period does not coincide with the main test of hypothesis H1, and many of the firms included in one period test are not included in the other test.

Table 5 presents the point estimates for 1963–1982. The results are similar to the 1972–1991 results (table 1). For example, β for current ratio is 42.5 percent in 1963–1982, 37.5 percent in 1972–1991; β for NOI to sales is 5 percent for the first period and 7.5 percent for the later period. These results indicate that β is stable over time.

VI. Conclusion

In this study we measure the rate of adjustment of selected financial ratios to equilibrium values. When firms experience a shock that affects their liquidity, equilibrating forces counterbalance a little more than a third of the shock in the next period. This finding suggests that firms' liquidity ratios have equilibrium values. Any movement away from the equilibrium value leads to a sharp adjustment to return to equilibrium. Performance measures, capital structure, and the gross margin ratio also show evi-

---

[29] Using a longer period would include data from a time when the firm was not contemplating merger. Using a shorter time period is problematic because there are insufficient observations to calculate meaningful correlation coefficients.

[30] Mann-Whitney U-tests showed similar results and are not reported.

Table 4
Test of Mean Adjustment Coefficients for Control Versus Bankrupt and Acquired Firms

	Number of Firms	Point Estimate of β	Significance Level Kolmogorov-Smirnov Two-Sample Test
Panel A:			
Current Ratio			
Acquisition or Merger	564	42.5%	not significant
Control	1,093	40.0%	
Quick Ratio			
Acquisition or Merger	546	42.5%	not significant
Control	945	42.5%	
Cash Plus Short-Term Investments/Current Assets			
Acquisition or Merger	554	47.5%	0.06
Control	944	42.5%	
Inventory/Current Assets			
Acquisition or Merger	507	45.0%	0.02
Control	894	37.5%	
Current Asset Decomposition			
Acquisition or Merger	498	50.0%	0.01
Control	1,262	45.0%	
Net Operating Income/Sales			
Acquisition or Merger	587	20.0%	0.02
Control	1,173	15.0%	
Net Operating Income/Assets			
Acquisition or Merger	588	20.0%	0.11
Control	1,173	17.5%	
Panel B:			
Primary EPS - Excluding Extraordinary Items			
Acquisition or Merger	643	32.5%	0.00
Control	1,896	45.0%	
Primary EPS - Including Extraordinary Items			
Acquisition or Merger	644	41.0%	0.00
Control	1,900	52.5%	
Fully Diluted EPS - Excluding Extraordinary Items			
Acquisition or Merger	516	32.5%	0.00
Control	1,625	40.0%	

Note: β measures the percentage of mean reversion of a ratio. The appendix provides detailed descriptions of the ratios.

dence of equilibrating forces, although the adjustment process is much longer. The adjustment speed of EPS ratios is quite high. About a third to half of the deviation from the equilibrium value is adjusted within one year. This result is consistent with Beaver (1970) and Ball and Watts (1972), who found that EPS ratios are mean reverting over several years. EPS including extraordinary items has a faster adjustment than EPS excluding extraordinary items.

Table 5

Point Estimates of Adjustment Coefficients, 1963–1982

		Point Estimate	
	Number of Firms	*Adjustment Coefficient* β	*Chi-Square*
Liquidity Measures			
Current Ratio	1,123	42.5%	12.0
Quick Ratio	952	35.0%	11.4
Cash Plus Short-Term Investments/Current Assets	964	37.5%	16.2
Inventory/Current Assets	927	35.0%	11.0
Current Asset Decomposition	907	45.0%	17.4
Performance Measures			
Net Operating Income/Sales	1,096	5.0%	21.8
Net Operating Income/Assets	1,096	2.5%	16.7
Earnings Per Share			
Primary - Excluding Extraordinary Items	746	20.0%	21.2
Primary - Including Extraordinary Items	604	35.0%	18.7
Fully Diluted - Excluding Extraordinary Items	148	40.0%	8.6
Capital Structure			
Equity/Debt	789	0.0%	23.6
Gross Margin Ratio	1,036	10.0%	17.2

Note: β measures the percentage of mean reversion of a ratio. The point estimate represents the minimum χ^2_q for all possible q values between zero and one at intervals of 2.5 percent. $\chi^2_q = 27.6$ at the 5 percent significance level with 17 degrees of freedom. The appendix provides detailed descriptions of the ratios.

The evidence from supplementary tests suggests that, for all ratios other than EPS, smaller firms adjust their ratios to the optimal target more swiftly than large firms. Separating the total adjustment effect into an industry and a mangement component shows that management's speed of adjustment is faster than that of industry. A comparison of the relative weight of the components suggest that both contribute a significant share to the total adjustment process.

Several tests were conducted to test the predictive power of the theoretical structure and the findings of the study. We found that control of NOI ratios over a 19-year period was correlated with the actual NOI ratio value in the twentieth year. This implies that the past value of the adjustment ratio is a predictor of future firm performance. Also, firms that subsequently were acquired showed better than average control of their liquidity and performance ratios. Firms that subsequently were acquired showed a destabilization of EPS ratios. Finally, the behavior of the adjustment over time was consistent.

Appendix

Ratio Name	*COMPUSTAT Items*
Current ratio	Item 4/Item 5
Quick ratio	(Items 1+2)/Item 5
Cash plus short-term investments/current assets	Item 1/Item 4
Inventory/current assets	Item 3/Item 4
Current asset decomposition	−(Item 1/Item 4)*log (Item 1/Item 4) −(Item 2/Item 4)*log (Item 2/Item 4) −(Item 3/Item 4)*log (Item 3/Item 4)
Net operating incomes/sales	(Item 13−Item 14)/Item 12
Net operating income/assets	(Item 13−Item 14)/Item 6
Primary earnings per share excluding extraordinary items and discontinued operations	Item 58
Primary earnings per share including extraordinary items and discontinued operations	Item 53
Fully diluted earnings per share excluding extraordinary items and discontinued operations	Item 57
Equity/debt	(Items 10+11)/(Items 5+9)
Gross margin ratio	(Item 12−Item 41)/Item 12
Equity/fixed assets	(Items 10+11)/Item 8
Sales/fixed assets	Item 12/Item 8
Sales/equity	Item 12/(Items 10+11)
Retained earnings/assets	Item 36/Item 6

References

Altman, E. I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23 (September): 589–609.

——, R. G. Haldeman, and P. Narayanan. 1977. Zeta analysis: A new model for identifying bankruptcy risk of corporations. Working paper No. 119, Salomon Brothers' Center for the Study of Financial Institutions, New York.

Ball, R., and R. Watts. 1972. Some time series properties of accounting income. *Journal of Finance* 27 (June): 663–82.

Beaver, W. H. 1966. Financial ratios as predictors of failure. *Journal of Accounting Research* 4 (Supplement): 71–111.

——. 1970. The time series behavior of earnings. *Journal of Accounting Research* 8 (Supplement) 62–99.

Brooks, L. D., and D. A. Buckmaster. 1976. Further evidence of the time series properties of accounting income. *Journal of Finance* 31 (December): 1359–73.

Brown, L. D., R. L. Hagerman, P. A. Griffen, and M. E. Zmijewski. 1987. Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting & Economics* 9 (April): 61–87.

Chesser, D. L. 1974. Predicting loan noncompliance. *Journal of Commercial Bank Lending* 56 (August): 28–38.

Collins, W. A., and W. S. Hopwood. 1980. A multivariate analysis of annual earnings forecasts generated from quarterly forecasts of financial analysts and univariate time-series models. *Journal of Accounting Research* 18 (Autumn): 390–407.

Cyert, R. M., and J. G. March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.

Davis, H. Z., and Y. C. Peles. 1992. On measuring serial correlation: Implications for finance and accounting research studies. *Review of Quantitative Finance and Accounting* 2 (September): 291–98.

Deakin, E. B. 1972. A discriminant analysis of predictors of business failures. *Journal of Accounting Research* 10 (Spring): 167–79.

Eckstein, O., and G. Fromm. 1968. The price equation. *The American Economic Review* 58 (December): 1159–83.

———, and D. Wyss. 1972. Industry price equation. In *The Econometrics of Price Determination*, edited by O. Eckstein. Washington, D.C.: Publications Services, Division of Administrative Services, Board of Governors, Federal Reserve System, 133–65.

Edmister, R. O. 1972. An empirical test of financial ratio analysis for small business failure prediction. *Journal of Financial and Quantitative Analysis* 7 (March): 1477–91.

Elam, R. 1975. The effect of lease data on the predictive ability of financial ratios. *The Accounting Review* 50 (January): 25–43.

Fama, E. F., and K. R. French. 1988. Permanent and temporary components of stock prices. *Journal of Political Economy* 96 (2): 246–73.

Foster, G. 1986. *Financial Statement Analysis*, 2d ed. Englewood Cliffs, NJ: Prentice-Hall.

Freeman, R. N., J. A. Ohlson, and S. H. Penman. 1982. Book rate-of-return and prediction of earnings changes: An empirical investigation. *Journal of Accounting Research* 20 (Autumn): 639–53.

Hall, R. L., and C. J. Hitch. 1939. Price theory and business behavior. *Oxford Economic Papers* 2 (May): 12–45.

Horrigan, J. O. 1965. Some empirical bases of financial ratio analysis. *The Accounting Review* 40 (July): 558–68.

Hotelling, H. 1953. New light on the correlation coefficient and its transforms. *Journal of the Royal Statistical Society* 15 (Series B): 193–232.

Kaplan, A. D. N., J. B. Dirlam, and R. F. Lanzillotti. 1958. *Pricing in Big Business—A Case Approach*. Washington, D.C.: Brookings.

Kendall, M., A. Stuart, and J. K. Ord. 1983. *The Advanced Theory of Statistics*, Vol. 3, 4th ed. London: Charles W. Griffin.

Lev, B. 1968. The aggregation problem in financial statements: An informational approach. *Journal of Accounting Research* 6 (Autumn): 247–61.

———. 1969. Industry averages as targets for financial ratios. *Journal of Accounting Research* 7 (Fall): 290–99.

Libby, R. 1975. Accounting ratios and the prediction of failure: Some empirical evidence. *Journal of Accounting Research* 13 (Spring): 150–61.

Miller, M. H. 1977. Debt and taxes. *Journal of Finance* 32 (May): 261–76.

Modigliani, F., and M. H. Miller. 1958. The cost of capital, corporation finance, and the theory of investment. *American Economic Review* 48 (June): 261–97.

Mood, A. M., F. A. Graybill, and D. C. Boes. 1974. *Introduction to the Theory of Statistics*, 3d ed. New York: McGraw-Hill.

Morck, R., A. Shleifer, and R. W. Vishny. 1988. Management ownership and market valuation. *Journal of Financial Economics* 20 (January/March): 293–315.

Peles, Y. C., and M. I. Schneller. 1989. The duration of the adjustment process of financial ratios. *The Review of Economics and Statistics* 62 (November): 527–32.

Ronen, J., and S. Sadan. 1981. *Smoothing Income Numbers: Objectives, Means, and Implications*. Reading, MA: Addison-Wesley.

Saxton, C. 1941. *Economics of Price Determination*. Oxford, England: Oxford University Press.

Silberston, A. 1970. Survey of applied economics: Price behavior of firms. *The Economic Journal* 80 (September): 511–82.

Wilder, R. P., C. G. Williams, and D. Singh. 1977. The price equation: A cross-sectional approach. *American Economic Review* 67 (September): 732–40.

Wood, A. 1975. *A Theory of Profits*. Cambridge, England: Cambridge University Press.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 748–764

Analyzing the Variance Investigation Decision: The Effects of Outcomes, Mental Accounting, and Framing

Marlys Gascho Lipe
University of Colorado

SYNOPSIS AND INTRODUCTION: In evaluating the variance investigation decision made by a manager, information *ex ante* or *ex post* to the decision may have an impact. Although normative models generally consider only *ex ante* data, research on hindsight or outcome bias has shown that performance evaluations are often affected by decision outcomes. This outcome effect can be explained through the following cognitive links. First, the outcome of the investigation will have an impact on the perceived benefits of the investigation. Second, as posited by decision research on mental accounting, investigation expenditures matched with perceived benefits are framed as costs while those without perceived benefits are framed as losses. Third, evaluators with a cost frame provide higher performance ratings than those with a loss frame.

A series of experiments with students and members of the Institute of Management Accountants demonstrates that managers making variance investigation decisions were evaluated more favorably when investigations revealed problems in the system. This is the well-known outcome effect. Further, the investigation outcome affected the perceived benefits from the investigation and, as predicted by mental accounting, expenditures with perceived benefits were framed as costs while those without perceived benefits were framed as losses. Finally, these frames affected performance evaluation as predicted. Thus, the outcome effect on performance evaluation can now be understood within a framework that includes the cognitive impact of mental accounting and framing.

My thanks to Richard Helleloid and to participants at the accounting workshops at the University of Oklahoma, Oklahoma State University, Pennsylvania State University, and the University of Texas at Arlington for helpful suggestions on an earlier draft. Steve Salterio and Andy Terner provided research assistance and the Ann Arbor chapter of the IMA provided experimental subjects for which I am grateful.

Submitted July 1992.
Accepted March 1993.

Surveys report that 85 percent of large U.S. manufacturing companies use standard costing systems (Cress and Pettijohn 1985; Gaumnitz and Kollaritsch 1988). Standard costing is particularly useful for managerial control purposes since variances from standards can be calculated for management use (Horngren and Foster 1987). At that point, managers must decide whether to collect more information regarding the performance reflected in the variances. Mathematical models have been formulated to analyze this variance investigation decision (e.g., see Kaplan 1969; Jacobs and Marshall 1984). Although these mathematical models provide normative or optimal prescriptions for the variance investigation decision, managers may not follow the recommendations of the models if their decisions are judged using criteria unrelated to the models (Magee 1976). Research on human information processing and decision making has shown that judgments are affected by outcomes (or hindsight) and mental accounting (a form of decision framing). If the managerial evaluations are affected by these phenomena, managerial actions will be effected by their potential impact.

This study investigates the effects of outcomes and framing on the evaluations of managers responsible for the variance investigation decision. Specifically, it develops a cognitive mechanism for the impact of outcomes on such evaluations using prior results and ideas from judgment and decision making research. Although the impact of hindsight and outcomes has been demonstrated in prior work in accounting (Helleloid 1988; Brown and Solomon 1987), a satisfactory cognitive explanation for the effect remains elusive (but, see Brown and Solomon 1991). Thus, the contribution of the present study is in the development and testing of this cognitive mechanism.

Key Words: *Variance investigation, Outcome effects, Hindsight, Mental accounting, Framing.*

Data Availability: *Data gathered and used in this study will be made available by the author upon request.*

THE remainder of this paper is organized as follows. The next section describes very simply the context of a variance investigation decision. A discussion of outcome and framing effects follows this description. Experimental work is then presented, followed by a discussion of results.

I. The Variance Investigation Decision

Most graduate and undergraduate introductory managerial accounting courses teach a simple cost/benefit model for use in making variance investigation decisions. The idea is that managers compare the expected costs under two alternative actions, where the variance is investigated or not investigated, and the action with the lower expected cost is prescribed. The expected costs included in the analysis are (1) an investigation expenditure (IE), which may be the cost of managerial time and/or testing procedures; (2) the cost of correcting any problems found in the system (bringing an "out of control" system back "in control"); and (3) the incremental costs of failing to correct an "out of control" or problematic system. Managers are assumed to estimate these costs and the associated probabilities (using past data or discussions with engineers, etc.) and then calculate the expected costs in order to arrive at a normative or optimal decision regarding investigation.

If the variance is investigated, IE is spent and the investigation will show whether the system is in or out of control. Thus, the manager gets feedback regarding his/her investigation decision if he/she investigates the variance.[1] With an investigation, two outcomes are possible: the system is in control (i.e., has no problems) and IE has been spent, or the system is out of control (i.e., has a problem) and IE has been spent. In either case, the expenditure and the results of the investigation will need to be reported to the next level of management, which will evaluate the performance of the decision maker (DM).

II. The Judgmental Impact of Outcomes

Since the investigation decision is made *ex ante* (i.e., before results of the investigation are known), the DM should be evaluated based on the *ex ante* information (Edwards 1984; Fischhoff 1983a) when that information is shared by the DM and the evaluator (Hershey and Baron 1992). Managers who choose actions with lower expected costs are likely to be evaluated more favorably than those who choose actions with higher expected costs. Indeed, Lipshitz (1989) found that those taking normatively appropriate actions were evaluated more favorably than others.

However, research on hindsight or "outcome bias" (Baron and Hershey 1988) shows that *ex post* information also affects performance evaluation. Studies show that it is difficult to ignore intervening events or outcomes in evaluating the performance of decision makers. For example, Brown and Solomon (1987) found that business students evaluated a capital budgeting committee's decision differently after a project's success or failure was reported than they did before the outcome data were available.[2] Similarly, Lipshitz (1989) showed that Israeli Defense Force officers evaluated a regiment commander more favorably when his decision to go to the aid of an attacked force was followed by a significant battle rather than a minor one. Furthermore, a commander who did not go to the attacked force's aid was viewed more favorably if this decision preceded an attack on his own sector (see also Mitchell and Kalb 1981; Baron and Hershey 1988). Each of these studies demonstrates that decision outcomes affect evaluations of the decision maker.[3]

Thus, in the case of a variance investigation, it is likely that evaluators will consider the outcome of the investigation (the *ex post* information), as well as the *ex ante* expected costs. Specifically, when the investigation reveals a problem, the manager's decision will appear more appropriate and will lead to higher performance ratings than when the investigation indicates no problem with the system. This leads to the following prediction:

---

[1] Such feedback is certainly less clear if he/she does not investigate.

[2] This effect of outcome information, though, was found to interact with the students' prior involvement in the decision task. Perhaps this involvement affected the amount of information actually shared by the DM and the evaluator in the task (Hershey and Baron 1992).

[3] Although the term "hindsight bias" is intuitively appealing for referring to this outcome effect, the literature has made a distinction between the impact of outcomes on remembered probabilities and on evaluations. The former is generally called the hindsight bias and the latter the outcome effect. The relation of the two can be imagined quite easily. For example, Fischhoff and Beyth (1975) posited that *ex post* data cause decision makers to revise estimates of prior probabilities which could lead to changes in the assessed optimality of the original decision. These changes, then, would likely affect the evaluation of the DM.

> H1: The outcome of a variance investigation will be related to the performance rating of the DM, such that his/her performance evaluation will be higher if the system was found *ex post* to be out of control than if it was found to be in control.

The management and organizational behavior literature indicates that a considerable number of motivational problems arise from the impact of *ex post* data on performance evaluations (for an introduction to its relation to agency theory, see Kaplan and Atkinson 1989). These theorists argue that it is important that employees believe their evaluations to be fair and to be dependent on their actions (Henderson 1985; Lawler 1981). They state that when the outcomes of managers' decisions are affected by events outside of the managers' control, these outcomes are "contaminated" measures of the managers' performance.

However, there are reasonable explanations for hindsight or outcome effects on performance evaluation (Brown and Solomon 1987). Therefore, these effects should not necessarily be labeled a "bias." In some cases, for example, outcomes may be considered in evaluating a manager in order to spread some risk to that manager. Risk-sharing agreements may be negotiated and contracted quite explicitly, and such use of outcome data should certainly not be considered a bias.

Although the impact of outcome information on performance evaluation is interesting and important, prior research has already demonstrated this effect in other settings. A more interesting issue for the present study is *how or why* outcome knowledge affects the evaluation. The contribution of the present study is in explicating how outcome knowledge affects the mental representation or decision frame of the evaluator, so as to affect performance evaluation.

III. The Judgmental Impact of Mental Accounts and Decision Framing

Researchers have posited the use of a psychological or mental account to evaluate actions with multiple consequences (Hirst et al. in press; Tversky and Kahneman 1981; see also Thaler 1990). Kahneman and Tversky (1984) indicate that this mental account is a way of framing a decision. The mental account is used to gather and combine the information that will be used in making a decision, while data considered irrelevant to the decision are relegated to a separate account. For example, Tversky and Kahneman (1981) asked subjects to evaluate the following situation:

> Imagine that you have decided to see a play where admission is $10 per ticket. As you enter the theater you discover that you have lost a $10 bill. Would you still pay $10 for a ticket for the play?

Eighty-eight percent of the subjects indicated a willingness to purchase the ticket. Another group of subjects was told that "you discover that you have lost the ticket. The seat was not marked and the ticket cannot be recovered. Would you pay $10 for another ticket?" In this situation, only 46 percent of the subjects were willing to buy another ticket. Tversky and Kahneman posit that, in the first case, the lost bill was not integrated into the same mental account with the $10 ticket cost. However, in the second case, the cost of the two tickets (the original and the replacement) were both posted to the same account so that the total cost became excessive.

Mowen and Mowen (1986) provided a more elegant research study of the effects of mental accounting (i.e., the use of mental accounts) or framing. They tested whether students' and business managers' propensity to take advantage of a constant-dollar trade discount would be affected by the size of the bill with which the discount was compared (i.e, the amount in the mental account). Consistent with prospect theory's S-shaped value function (Kahneman and Tversky 1979), both groups were more likely to pursue the discount when the order size was small.

Although the reviewed research has shown that a person's frame or mental account can affect decision making (see also Thaler 1980), there has been little accounting research on framing effects. Perhaps this is because most studies of judgment and decision making in accounting have focused on auditors for whom the decision frame may be fairly well specified (but see Ashton et al. 1988).[4] However, Shields et al. (1987) tested whether auditors' judgments would differ when asked to provide an estimate of audit values for some accounts versus when asked to provide estimates of book value misstatements for those accounts. Although final audit values did not differ under the two "frames," there was no *a priori* theory predicting such a difference. This study, in contrast, takes an approach similar to Mowen and Mowen (1986), using prior decision research to make predictions about the differences in judgmental frames.

Framing Effects in Variance Investigation

Two decision frames seem to be possible in the variance investigation situation. The investigation expenditure, IE, can be framed as a cost or as a loss. Kahneman and Tversky (1984) suggest that the frame to be invoked will depend on whether there is a benefit or "advantage" to match with the expenditure (i.e., to place in the same mental account). Thus, IE will be considered a cost if some benefit is also posted to the psychological account containing the expenditure. However, if no benefit is perceived to relate to that account, IE will be considered a loss. As an oft-cited example of the impact of a benefit posted to a psychological account, Schoemaker and Kunreuther (1979) showed that presenting a choice between a sure loss and a probabilistic loss as (1) an insurance decision or as (2) a gambling decision led to significantly different choices. Although the monetary effects are the same, people displayed a greater propensity to take the sure loss (i.e., pay the premium) in the insurance context than in the gambling context. Researchers have suggested that this may be true because the term "insurance" implies that a benefit is received for the money (Schoemaker 1980, 83; Slovic et al. 1982) so that the expenditure is considered a cost when described as insurance, but considered a loss when described as a gamble.

Whether a benefit is perceived to result from a variance investigation will likely be affected by the investigation's outcome. Thus, if the system is found to be out of control (and therefore corrected), the firm would have received a benefit in the form of a corrected system.[5] However, if an investigation reveals that the system is already performing correctly, an evaluator may feel that nothing was gained by the search (see section II on the effect of hindsight). This relation is posited in hypothesis H2:

---

[4] However, some studies have considered auditors' hypothesis-testing strategies, specifically whether a confirmatory or disconfirmatory "frame" was used (e.g., Trotman and Sng 1989).

[5] Further, Thaler and Johnson (1990) and Linville and Fischer (1991) showed that people prefer to integrate (or combine) outcomes in situations where a small loss or expenditure (such as IE) is coupled with a large gain (such as the gain from bringing an out-of-control system back to in-control). Thus, the use of a mental account seems particularly plausible in evaluating a variance investigation decision.

> H2: The outcome of an investigation will be related to the perceived benefit of the investigation, such that a greater benefit from the investigation will be perceived when the the system is found to be out of control than when the system is found to be in control.

Although the idea that a benefit implies a cost (while expenditures without benefit are losses) has been presented in many discussions of mental accounting (Tversky and Kahneman 1981; Kahneman and Tversky 1984; Thaler 1985), it has not been tested. Since predicting decision frames has proven to be very difficult in some contexts (Fischhoff 1983b), it would be particularly interesting to test this idea which is stated in hypothesis H3:

> H3: Perceived benefits from an investigation will be related to the framing of IE, such that IE would be considered to be a cost to the extent that a benefit is perceived to result from it, and is considered to be a loss to the extent that no such benefit is perceived to result.

Since research has shown that a person's frame or mental representation of a decision can affect his/her judgment (Kahneman and Tversky 1984; Mowen and Mowen 1986; Thaler 1980), a relation is expected between an evaluator's frame and the evaluation itself. Thus, if the perception of a benefit leads to framing IE as a cost, a relatively positive evaluation of the manager who made such a payment is expected. In contrast, the manager who is held responsible for a loss (i.e., a payment without perceived benefit) is likely to be evaluated relatively lower. The following hypothesis results:

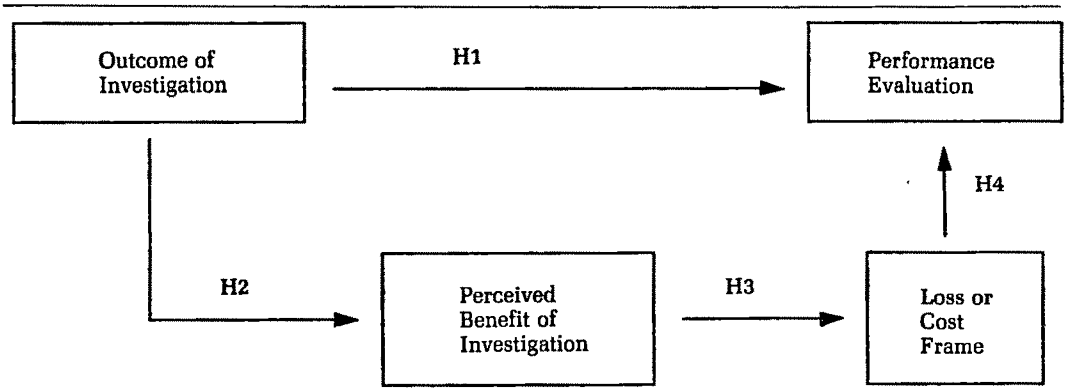
> H4: An evaluator's decision frame will be related to the performance rating of the DM such that a manager is rated higher when IE is framed as a cost than when IE is framed as a loss.

Thus, experimental work related to mental accounts and decision framing will test whether outcomes of a variance investigation affect perceived benefits, whether these perceived benefits predict the loss and cost frames, and whether the frames are related to performance evaluations of managers making the investigation decision. Figure 1 presents a schematic representation of each hypothesis' predicted relations and effects. This figure shows that mental accounting and decision framing may provide an explanation for the outcome effect. Hypothesis H1 predicts this link of outcome and performance evaluation. The remaining hypotheses, H2, H3, and H4, then explicate the cognitive mechanisms underlying this link, indicating *how and why* outcomes affect performance evaluations in this context.

IV. Research Method

Three experiments will be reported: the first two use a common scenario or context and common procedures that can be briefly described as follows. Subjects were students in an undergraduate cost accounting course who had just studied variance analysis and the variance investigation expected cost model as described in section I. The students were presented with a variance investigation decision (i.e., data for IE and the other relevant costs and probabilities) and were asked to make the decision (to be sure they understood the normative model). The students were then told the manager chose

Figure 1
Hypothesized Relations and Effect



to investigate the variance and were provided with information regarding the outcome of the investigation. Next, they were asked to evaluate the manager's performance using a rating scale, and were asked to provide some judgments regarding the frame they used to evaluate the task.

The first experiment tests the effects of *ex post* information as posited in hypothesis H1. It also tests the predictions of hypotheses H2 and H3 regarding the decision frames invoked in the variance investigation task and hypothesis H4 regarding the effect of these frames on the performance evaluation.

Experiment One

Subjects, Design, and Procedure: One hundred and forty-two students in a cost accounting course participated in the experiment as a part of the class. Students received a small amount of course credit for the task. The task, administered at the end of a lecture on variance investigation, took about five minutes.

This experiment had a 2 x 2 between-subjects design, where the variables were normative decision (NORM) and investigation outcome (OUTCOME). The data presented to the subjects indicated that the normative decision was (1) to investigate or (2) not to investigate (using the simple expected costs model as our normative model). However, as this design variable was unrelated to the hypotheses of this study, it will not be discussed further.[6]

Subjects were presented with the following data related to a variance investigation decision:

> Assume that you are VP-Manufacturing for a medium sized firm. Your firm uses a standard costing system with currently attainable standards. You have several production managers working for you. Each one runs his/her own production plant.

---

[6] NORM was included here to relate the experimental task to the students' course material, where they were learning to use expected costs for making the investigation decision. Certainly, since management theories cited in section II state that DMs should be evaluated on *ex ante* information, one would expect investigating managers to be evaluated more highly when they are acting normatively than when they are acting contrary to a normative prescription. NORM's significant main effect (shown in table 1) bears out this prediction. NORM is included in all experiments (although operationalized differently in experiment three) to increase comparability.

> Pat Smith is one of those production managers. Below is information regarding a variance investigation decision that Pat faced this year.
>
> Pat's plant had a large direct labor efficiency variance this year. Pat had to decide whether to investigate the variance further. The cost of such an investigation (IE) was estimated at $4,000. Other estimates stated that the cost of correcting a problem would be $12,000 and the present value of future costs of an uncorrected out-of-control system would be $22,000. Past data show that the probability that the system is out of control (given this large reported variance) is 50 percent.

The normative decision for the above version is to investigate the variance. Students were asked to show the normative decision to be sure that they understood the expected cost model. Any student making the wrong calculation or judgment was excluded from the analyses.[7]

Students were then informed that Pat made the investigation. Some were told that the system was found to be in control, while others were told that it was out of control. Both groups were asked to make the following judgment:

> You, as VP-Manufacturing, received a report from Pat detailing all of the costs and estimates involved in the investigation decision, along with Pat's decision.
>
> You need to evaluate Pat's performance for the period. Please give a judgment or opinion on the scale below. Although Pat is not held accountable for direct material or overhead variances, the direct labor variances and investigation decisions are Pat's responsibility.
>
> Pat's performance this period was (make a slash on the scale below):
>
> 0 50 100
> |----|----|----|----|----|----|----|----|----|----|
> Very Below Average Above Very
> Poor Average Average Good

One-half of the experimental subjects were then asked to describe the frame they used to evaluate the investigation expenditure.[8] Specifically, they were presented with four frames (listed in figure 2) and asked to indicate which of the frames best described their own thinking about the task. These frames provide a comment about IE's benefit to the company (either Benefit or No benefit) and a comment regarding the IE's loss- or cost-like nature (Loss or Cost). For example, frame BL indicates that a benefit has accrued but that IE is a loss.[9] These frame choices are used to test hypotheses H2, H3, and H4.

Results

Performance Evaluations (Hypothesis H1). The performance evaluations were analyzed using a 2 x 2 analysis of variance (ANOVA); the results are presented in panel A of

---

[7] Only 24 out of 282 students (8.5 percent) participating in all experiments made incorrect calculations.

[8] The remainder of experiment one's students performed a follow-up task similar to experiment two. Although the results of tests of hypotheses H3 and H4 for these subjects were the same as the results reported for experiment two, results for a test of hypothesis H2 were not. However, it seems quite possible that the insensitivity of the dependent variable used to measure benefit perception may have contributed to the insignificant results here. Students were asked to indicate whether they perceived a benefit by checking off a box stating "yes," "no," or "no opinion." Only six of the 70 participants indicated no benefit. Therefore, although the framing and mental accounting effects were found for these subjects (hyp. H3 and H4), it was decided that a response *scale* could more sensitively measure benefit perception and could allow a more powerful test of hypothesis H2. This methodology and test are reported in experiment two.

[9] Six different orders of the four frames were constructed. A chi-squared test indicates that order did not affect the chosen frame ($\chi^2 = 10.08$, $df = 15$, $p = 0.81$). The subjects had to make a forced choice as there was no option for choosing none of the frames.

Figure 2
Frames Presented in Experiment One

Frame	*Description*
NL	The firm spent $4,000 on a variance investigation. The firm got nothing in return for that money. There was no benefit from spending that. It seems like that $4,000 was a loss of resources for the company.
NC	The firm spent $4,000 on a variance investigation. The firm got nothing in return for that money. There was no benefit from spending that. The $4,000 was just another cost of doing business for the company.
BC	The firm spent $4,000 on a variance investigation. The firm got something in return for that money. There was real benefit from spending that. The $4,000 was just another cost of doing business for the company.
BL	The firm spent $4,000 on a variance investigation. The firm got something in return for that money. There was real benefit from spending that. It seems like that $4,000 was a loss of resources for the company.

table 1. OUTCOME had a statistically significant main effect on the performance ratings ($F=9.19$, $p<0.01$). These performance ratings, presented in panel B of table 1, indicate that investigating managers were viewed more favorably when a problem was found in the system than when the system was in control. This is the well-known outcome effect posited in hypothesis H1.

Frame Choices and Effects (*Hypotheses H*2, *H*3, *H*4). As the outcome effect is present in this context, hypotheses H2, H3, and H4 posit the mechanisms by which this effect occurs. Tests of these hypotheses are more process-oriented than tests of hypothesis H1, in that subjects were asked how they thought about the task. Specifically here, they were asked to choose among the frames listed in figure 2.

Hypothesis H2 posits that the outcome of the investigation will affect the perceived benefit from the investigation. A chi-square test of the relevant data indicates a relation of investigation outcome and perceived benefit ($\chi^2=15.13$, $df=1$, $p<0.01$). Specifically, 33 of 36 students who were told that the system was found to be out of control indicated that there was a benefit from investigating the variance (i.e., chose frame BC or BL). In contrast, only 18 of 36 students who were told that the system was found to be in control perceived such a benefit.

Hypothesis H3 implies that students' frame choices would generally fall into BC and NL, as expenditures with benefits are costs of business (BC) while such payments without benefits are losses (NL). Of the 72 students, 55 chose either BC or NL (76 percent). A chi-square test indicates a relation between perceived benefit and loss/cost judgment ($\chi^2=14.82$, $df=1$, $p<0.01$) as predicted in hypothesis H3. Specifically, 21 students indicated that no benefit accrued from the investigation and 14 of these 21 also indicated the payment was a loss rather than a cost. In comparison, 41 of the 51 students indicating a benefit felt IE to be a cost rather than a loss.

Experiment one also provides a test of hypothesis H4. An F-test indicates that performance evaluations differed for students using the different frames ($F=7.31$, $df=3$, 68, $p<0.01$). Mean performance evaluations for each frame were 30.64 (NL), 49.29

Table 1
Experiment One Performance Ratings

Panel A. Results of the Analysis of Variance:

Source	*SS*	*df*	*MS*	*F-stat*	*Signif.*
NORM (N)	9,304.93	1	9,304.93	19.80	0.00
OUTCOME (O)	4,319.45	1	4,319.45	9.19	0.00
N x O	67.61	1	67.61	0.14	0.71
Error	64,848.88	138	469.92		
Total	78,540.87	141			

Panel B. Mean Performance Ratings:

Normative Decision	*In-Control System*			*Out-of-Control System*			
	Mean	*Std. Dev.*	*n*	*Mean*	*Std. Dev.*	*n*	*Mean*
Investigate	56.75	22.36	36	66.41	21.28	37	61.64
Do Not Investigate	39.17	20.51	35	51.59	22.51	34	45.29
Mean	48.08			59.31			

(NC), 51.60 (BL), and 61.61 (BC). Further, the mean evaluations for students indicating a cost and a loss frame, respectively, were 59.81 and 39.37, a statistically significant difference ($t = 3.66$, $p < 0.01$).

Experiment Two

Results of experiment one supported all of the hypotheses. However, the subjects' choices were quite constrained in the frame elicitation task used to test hypotheses H2, H3, and H4. Borrowing a methodology from Fischhoff (1983b), experiment one asked the subjects to choose among four frames which were briefly described. Because forced-choice tasks may provide experimental results through experimenter demand and a lack of freedom in responses, a second experiment was developed to provide an additional test of these hypotheses using a different methodology. Here, rating scales were utilized to assess the subjects' frames, placing fewer constraints on their responses.

Subjects, Design, and Procedure. A new group of 116 cost accounting students participated in the experiment following a lecture on variance investigation. The experimental design was identical to that in experiment one; NORM and OUTCOME were manipulated between subjects. The students were asked to calculate the correct decision for a variance investigation task, were told the manager had chosen to investigate, were given the outcome of the investigation, and were asked to evaluate the manager's performance.

After completing these tasks, the students rated the investigation expenditure along two dimensions: the degree to which it provided a benefit to the firm, and the degree to which it was a loss or a cost to the firm. The questions and rating scales were presented in the following form (the order of these questions was varied and this order was found to have a statistically insignificant effect):

> In the situation I presented to you, the firm spent $4,000 on a variance investigation. Do you feel that the firm got something in return for that money? Did the $4,000 expenditure result in some benefit to the company? Please make a slash on the scale below to indicate your opinion.

0	50	100
|----|----|----|----|----|----|----|----|----|----|		
No, there was no benefit	No Opinion	Yes, there was a benefit

> In the situation I presented to you, the firm spent $4,000 on a variance investigation. When you think about that $4,000, do you tend to think of it as a COST or as a LOSS to the company? I just want your opinion. How do you tend to think about it? Please make a slash on the scale below to indicate your opinion.

0	50	100
|----|----|----|----|----|----|----|----|----|----|		
LOSS	No Opinion	COST

Results

Hypothesis H2 predicts that the outcome of the investigation will affect the perceived benefit. The mean benefit rating for the "in control" group was 71.22 while that for the "out of control" group was 84.95. These differed significantly ($t=3.80$, $p<0.01$);[10] again providing support for hypothesis H2.

Further, hypotheses H3 and H4 are also supported. The correlation of perceived benefit and loss/cost rating, as predicted by hypothesis H3, was significant across all subjects ($r=0.61$, $df=114$, $p<0.01$). Categorizing responses as "no benefit" or "loss" when below 50 and "benefit" or "cost" when above 50 allows a chi-square test of the relation between perceived benefit and frame. This test also provides support for hypothesis H3 (Yates corrected $\chi^2=18.31$, $df=1$, $p<0.01$).

Finally, the experiment supports hypothesis H4's predicted relation between performance judgments and decision frames. The correlation of performance judgments and judgments of loss/cost frame was 0.48 in this experiment, a statistically significant relation ($p<0.01$), and the mean evaluation for students indicating a cost frame (51 to 100 on the scale) and a loss frame (0 to 49 on the scale) were 59.94 and 34.17, respectively, a statistically significant difference ($t=6.30$, $p<0.01$).

Experiment Three

Experiments one and two yielded convergent results, showing that outcome information was related to the perceived benefits of an investigation (hyp. H2), perceived benefits were related to the loss or cost frame (hyp. H3), and this frame was related to the performance evaluation (hyp. H4). However, since student subjects were used in these experiments, generalizability to business managers may be questioned. Therefore, experiment three tests the findings of experiment two using business people as subjects.

---

[10] Levene's test for unequal variance showed a significant difference across the groups ($F=9.69$, $p<0.01$), so the reported t-test used unpooled variance.

Subjects, Design, and Procedure. The local chapter of the IMA (formerly the National Association of Accountants) was asked to assist in an academic study. Leaders of the group agreed to participate if the study could be run by mail and if the materials could be compressed to one sheet (two-sided) of paper. The stimuli material was included in a mailing of the monthly newsletter sent to 350 members. A postage-paid return envelope was also provided and the responses were returned anonymously to the researcher. Fifty-nine responses were received within one month.[11]

The design was the same as that used in experiment two; the normative decision (NORM) and the outcome of the investigation (OUTCOME) were manipulated in a 2 x 2 between-subjects design. Subjects evaluated the manager's performance, and then rated IE as to whether it provided a benefit to the firm and whether it was a loss or a cost to the firm.

Operationalizing the manipulated variables was more difficult for this experiment.[12] Relevant to the tested hypotheses, the OUTCOME variable was operationalized as follows:

> During the investigation, Pat found that the variance had been caused by a unique and temporary event. As it is very unlikely to happen again, no correction was needed. (In-control condition.)
>
> During the investigation, Pat found that the variance had been caused by a problem in the production system. This problem was not temporary and would continue until corrected. Pat, therefore, corrected the identified problem. (Out-of-control condition.)

Subjects were also asked their job title, the type of firm for which they worked, and how much full-time business experience they had. Finally, because it was expected that people's susceptibility to outcome effects in performance evaluation may be related to the evaluation system under which they work, the subjects were asked the following questions:

> When you are evaluated at work, are you held responsible for events and circumstances that are outside of your actual control?
>
> Do you feel it is appropriate, in your firm, to hold people responsible for events and circumstances outside of their actual control?

These will be referred to as the responsibility and appropriateness questions, respectively. Each question was followed by a 0 to 100 scale, as shown below:

0 50 100
|----|----|----|----|----|----|----|----|----|----|
never occasionally often all the time

Results

Debriefing Questions. The IMA members had an average of 15.18 years of full-time business experience, with a range of 1 to 40 years. Job titles were varied but the most

---

[11] Certainly, relative to the total chapter membership, this response rate (17 percent) is very low. However, the number of responses is quite close to the number of active chapter members. Average attendance at chapter technical meetings is approximately 50. The author made appeals for responses at two such meetings.

[12] Because few business people could be expected to recall or use the same normative model, a simplified normative model was used. Subjects were told that the firm's policy stated that "variances greater than 7 percent of standard costs allowed should be investigated by the manager (while those below this should not)." The variance was then described as 8 percent or 6 percent of standard costs allowed, indicating that the investigation should or should not take place.

Table 2
Experiment Three Performance Ratings

Panel A. Results of the Analysis of Variance:

Source	*SS*	*df*	*MS*	*F-stat*	*Signif.*
NORM (N)	1,134.64	1	1,134.64	4.88	0.03
OUTCOME (O)	11,164.55	1	11,164.55	48.01	0.00
N x O	1,145.31	1	1,145.31	4.93	0.03
Error	12,790.12	55	232.55		
Total	26,234.62	58			

Panel B. Mean Performance Ratings:

Normative Decision	*In-Control System*			*Out-of-Control System*			
	Mean	*Std. Dev.*	*n*	*Mean*	*Std. Dev.*	*n*	*Mean*
Investigate	54.89	17.96	18	73.83	14.21	12	62.47
Do Not Investigate	37.08	18.25	13	73.88	8.53	16	57.38
Mean	47.42			73.86			

frequent were controller or assistant controller (13), financial manager or VP-finance (8), and auditor (7). Another 11 subjects were cost accountants, general accountants, or accounting supervisors. Thirty of the 59 subjects were employed in manufacturing firms.

The average responsibility and appropriateness ratings did not differ across the experimental conditions. The average responsibility rating was 44.37, indicating that these professionals were sometimes held responsible for events outside of their control. However, the subjects felt that this was only "occasionally" appropriate in their firm (average appropriateness rating of 25.66). Since these people apparently believe that outcome effects are often inappropriate, it will be interesting to test whether their evaluations are affected by such outcomes.

Analysis of Outcome Data on Performance Ratings. The performance ratings were analyzed via a 2 x 2 ANOVA; the results are reported in panel A of table 2. As predicted by hypothesis H1, OUTCOME had a significant main effect ($F=48.01$, $p<0.01$). However, a significant interaction for OUTCOME and NORM ($F=4.93$, $p=0.03$) indicates that simple main effects must be examined. Panel B of table 2 presents the mean ratings for each cell of the experiment, indicating that the impact of OUTCOME was significant whether the investigation was prescribed by company policy ($t=2.70$, $p<0.01$), or not ($t=4.29$, $p<0.01$). Thus, the outcome effect (hyp. H1) was also present for this subject group.[13]

[13] In contrast to student subjects, here the impact of NORM was only felt when the ex post data showed the system to be in control ($t=2.46$, $p=0.01$). Note, however, that here the "normative" model is actually a company policy or rule of thumb rather than an optimal decision model.

Analysis of Frames. Experiment three reinforces the prior findings which support hypotheses H2, H3, and H4. The mean benefit rating for the "in-control" group was 63.29 while that for the "out-of-control" group was 92.21, a significant difference ($t=4.57$, $p<0.01$).[14] This result supports the prediction of hypothesis H2 that the outcome of the investigation will affect the perceived benefit. The correlation of perceived benefit and loss/cost rating across all subjects was significant ($r=0.47$, $z=3.56$, $p<0.01$), as predicted by hypothesis H3. Again, categorizing responses relative to their position above or below 50 on the scale and then using a chi-square test demonstrates that the perception of benefit and the loss/cost frames are not independent (Yates corrected $\chi^2=8.16$, $df=1$, $p<0.01$).

In experiment three the correlation of performance judgments and assessments of the loss/cost frame was 0.37, a statistically significant relation ($z=2.85$, $p<0.01$), supporting hypothesis H4. The mean evaluation for subjects indicating a cost frame (51 to 100 on the scale) and for those indicating a loss frame (0 to 49 on the scale) were 63.61 and 41.09, respectively, a statistically significant difference ($t=3.51$, $p<0.01$).

V. Discussion and Extensions

The results of these experiments provide insights into both performance evaluations and the prediction of frames. These two areas will be discussed separately, followed by suggestions for extensions of this research.

Managerial Performance Evaluation

The experimental results indicate that a manager's performance evaluation is affected by *ex post* information and the decision frame of the evaluator (see hyp. H1 and H4 in fig. 1). The importance and appropriateness of outcome effects on performance evaluation can only be evaluated within the environment of the particular decision maker and evaluator. The experimental task used in this study abstracted from the contracting and incentive environment of the firm in order to take a first look at the issue.[15]

Although hindsight and outcome effects are not necessarily a bias within a given environment, their existence can affect decisions. A rational manager who believes his/her evaluations will be affected by outcomes has reason to supplement the basic variance investigation decision model (or company policy) with these expected outcomes and personal payoffs or even to substitute them for the parameters of the basic model (Magee 1976; Prakash and Rappaport 1977).

It would be interesting to see some modeling of the investigation decision which includes these expected *ex post* data, outcome effects, and their costs and benefits. Brown and Solomon (1991) suggest that outcome effects may affect managers' abilities to learn from feedback and may be best dealt with through the design of the evaluation system (rather than by trying to train managers to avoid the effect).

---

[14] Levene's test for unequal variance again showed a significant difference across the groups ($F=31.74$, $p<0.01$), so the reported *t*-test used unpooled variance.

[15] Replies to the responsibility question in experiment three showed that firms vary considerably in how much responsibility they assign for uncontrollable factors. The range of responses was from 0 to 94, with a standard deviation of 27.09.

Implications for Framing Research

Fischhoff (1983b, 103) stated that his framing prediction "results were generally discouraging for the prediction of individuals' choices, generally encouraging for the prediction of group choices." In the current study, frames could be predicted based on perceived benefits of an expenditure. Framing researchers may need to explicate more carefully the mechanisms by which frames are invoked. By knowing more about these mechanisms and links, framing research can become both more predictive and more precise.

The current study also finds a relation between frames and performance evaluation. This result contrasts with Fischhoff's (1983b) difficulty in relating decision frames and decision choices, even though experiment one used a methodology very similar to Fischhoff's. He presented subjects with short explanations of different frames and asked them to indicate the one that seemed most natural or the one that they had used to think about the choice. However, subjects' preferences in making the actual choices were unrelated to their indicated frames. In experiment one of the current study, subjects also indicated the frame they used to think about the investigation expenditure, and their frames apparently did affect the performance evaluations they made of the manager.

Although it is difficult to determine why the two studies have different results, Fischhoff's subjects may have had less motivation to appear consistent than did the students and professionals in the current study. Students in the current study may have felt they needed to make consistent frame choices and evaluations since they were returning the materials to their instructor, and the business people may have felt that consistency was important to their perceived professionalism (or they simply may make more consistent judgments as a result of their training). In contrast, Fischhoff's subjects were recruited from a more heterogenous population, including students from throughout the university and people at an employment agency. Future research on framing's impact on judgments and decisions should consider the effect of motivation and justification (Ashton 1990; Hagafors and Brehmer 1983; Libby and Lipe 1992).

Extensions of the Study

Interesting extensions of this study include testing the proposed cognitive mechanisms for outcome effects in other contexts or tasks where prescriptive models or policies are used. Other such tasks include order or production quantities (EOQ), capital budgeting decisions, product choices, and sampling in an audit setting. Also, comparisons with Fischhoff's work suggest that the relation of decision frames with evaluations or decisions may be affected by accountability or reputation-building. Thus, future research on frame elicitation and decision making could test these effects directly.

Another fascinating question is how the cognitive framework developed here relates to other explanations offered for hindsight or outcome effects. Although no well-developed theory has been advanced for outcome effects, some suggestions have been offered. Brown and Solomon (1991) summarized these in three categories: those using a cognitive reconstruction explanation, those with a self-enhancing motive explanation, and those with an escalation of commitment explanation. While an analysis of these explanations as they relate to the framework developed here is beyond the scope of this article, designing tests which could discriminate among them or explore their relation to the framework would be a very important extension of this work. Also, Brown and Solomon (1987) showed that prior involvement in a decision task can reduce the outcome effect. An interesting integration of that work with the present study would test the impact of prior involvement on each of the cognitive links posited here to determine which of the links are affected by such involvement. Thus, the present paper has provided a cognitive explanation for the outcome effect and this should provide direction and impetus for future work in this area.

References

Ashton, R. 1990. Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback, and justification. *Journal of Accounting Research* 28 (Supplement): 148–80.

———, D. Kleinmuntz, J. Sullivan, and L. Tomassini. 1988. Audit decision making. In *Research Opportunities in Auditing: The Second Decade*, edited by A. R. Abdel-khalik and I. Solomon. Sarasota, FL: AAA.

Baron, J., and J. Hershey. 1988. Outcome bias in decision evaluation. *Journal of Personality and Social Psychology* 54: 569–79.

Brown, C., and I. Solomon. 1987. Effects of outcome information on evaluations of managerial decisions. *The Accounting Review* 62 (July): 564–77.

———, and ———. 1991. An experimental investigation of explanations for outcome effects on appraisals of capital-budgeting decisions. Paper presented at the Contemporary Accounting Research Conference.

Cress, W., and J. Pettijohn. 1985. A survey of budget-related planning and control policies and procedures. *Journal of Accounting Education* 3 (Fall): 61–78.

Edwards, W. 1984. How to make good decisions. *Acta Psychologica* 56: 5–27.

Fischhoff, B. 1983a. Decision analysis: Clinical art or clinical science? In *Human Decision Making*, edited by L. Sjoberg, T. Tyszka, and J. A. Wise. Bodfors, Sweden: Doxa.

———. 1983b. Predicting frames. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9: 103–16.

———, and R. Beyth. 1975. "I knew it would happen"—Remembered probabilities of once-future things. *Organizational Behavior and Human Performance* 13: 1–16.

Gaumnitz, B., and F. Kollaritsch. 1988. Manufacturing variances: Current practices and trends. Working paper, Ohio State University, Columbus, OH (May).

Hagafors, R., and B. Brehmer. 1983. Does having to justify one's judgments change the nature of the judgment process? *Organizational Behavior and Human Performance* 31: 223–32.

Helleloid, R. 1988. Hindsight judgments about taxpayers' expectations. *Journal of the American Taxation Association* 9: 31–46.

Henderson, R. 1985. *Compensation Management: Rewarding Performance*, 4th ed. Reston, VA: Prentice Hall.

Hershey, J. C., and J. Baron. 1992. Judgment by outcomes: When is it justified? *Organizational Behavior and Human Decision Processes* 53: 89–93.

Hirst, D. E., E. Joyce, and M. Schadewald. In press. Mental accounting and outcome contiguity in consumer borrowing decisions. *Organizational Behavior and Human Decision Processes.*

Horngren, C., and G. Foster. 1987. *Cost Accounting: A Managerial Emphasis*, 6th ed. Englewood Cliffs, NJ: Prentice Hall.

Jacobs, F., and R. Marshall. 1984. A note on the choice structure of cost variance investigation models. *Journal of Accounting Literature* 3: 73–83.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (March): 263–91.

———, and ———. 1984. Choices, values, and frames. *American Psychologist* 39: 341–50.

Kaplan, R. 1969. Optimal investigation strategies with imperfect information. *Journal of Accounting Research* 7 (Spring): 32–43.

———, and A. Atkinson. 1989. *Advanced Management Accounting*. Englewood Cliffs, NJ: Prentice Hall.

Lawler, E. 1981. *Pay and Organization Development*. Reading, MA: Addison-Wesley.

Libby, R., and M. Lipe. 1992. Incentives, effort, and the cognitive processes involved in accounting-related judgments. *Journal of Accounting Research* 30 (Fall): 249–73.

Linville, P., and G. Fischer. 1991. Preferences for separating or combining events. *Journal of Personality and Social Psychology* 59: 5–23.

Lipshitz, R. 1989. "Either a medal or a corporal": The effects of success and failure on the evaluation of decision making and decision makers. *Organizational Behavior and Human Decision Processes* 44: 380–95.

Magee, R. 1976. A simulation analysis of alternative cost variance investigation models. *The Accounting Review* 51 (July): 529–44.

Mitchell, T., and L. Kalb. 1981. Effects of outcome knowledge and outcome valence on supervisors' evaluations. *Journal of Applied Psychology* 66: 604–12.

Mowen, M., and J. Mowen. 1986. An empirical examination of the biasing effects of framing on business decisions. *Decision Sciences* 17: 596–602.

Prakash, P., and A. Rappaport. 1977. Information inductance and its significance for accounting. *Accounting, Organizations, and Society* 2: 29–38.

Schoemaker, P. 1980. *Experiments on Decisions Under Risk: The Expected Utility Hypothesis*. Boston, MA: Marinus Nijhoff.

———, and H. Kunreuther. 1979. An experimental study of insurance decisions. *Journal of Risk and Insurance* 46: 603–18.

Shields, M, I. Solomon, and W. Waller. 1987. Effects of alternative sample space representations on the accuracy of auditors' uncertainty judgments. *Accounting, Organizations, and Society* 12: 375–85.

Slovic, P., B. Fischhoff, and S. Lichtenstein. 1982. Response mode, framing, and information-processing effects in risk assessment. In *Question Framing and Response Consistency* edited by R. Hogarth. San Francisco, CA: Jossey-Bass.

Thaler, R. 1980. Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization* 1: 39–60.

———. 1985. Mental accounting and consumer choice. *Marketing Science* 4 (Summer): 199–214.

———. 1990. Anomalies: Saving, fungibility, and mental accounts. *Journal of Economic Perspectives* 4 (Winter): 193–205.

———, and E. Johnson. 1990. Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science* 36 (June): 643–60.

Trotman, K., and J. Sng. 1989. The effect of hypothesis framing, prior expectations and cue diagnosticity on auditors' information choice. *Accounting, Organizations, and Society* 14: 565–76.

Tversky, A., and D. Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211 (January): 453–58.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 765–782

Stewardship Value of "Distorted" Accounting Disclosures

Robert M. Bushman
Raffi J. Indjejikian
University of Chicago

SYNOPSIS AND INTRODUCTION: An issue of fundamental importance to accountants concerns the qualities possessed by, or that should be possessed by, accounting information. The usual intuition is that any distortion of accounting numbers relative to the true underlying cash flows of the firm is generally undesirable from the standpoint of investors. Indeed, among the primary qualitative characteristics of accounting information championed by the FASB in their conceptual framework is *reliability*, which is defined as "the quality of information that gives assurance that it is reasonably free of *error* and *bias* and is a *faithful representation*" (emphasis added). Normative statements about the desirable qualities of accounting information are often problematic since such information typically serves multiple purposes. For example, Gjesdal (1981) identifies a decision-making role and a stewardship role for accounting information and shows that the two needs may not be served equally well. The objective of this paper is to demonstrate that "distorted" accounting information may actually be preferred if the focus is on the stewardship value of accounting information.

"Distorted" accounting information is characterized as information signals that are *biased* relative to the expected value of the firm or that measure the value of the firm with *error* (white noise). Bias and noise are meant to be representative of the many inadequacies usually attributed to accounting information. Current Generally Accepted Accounting Principles prevent or delay the recognition of certain assets and liabilities and their income statement counterparts, generating what may be thought of as bias. For example, the asymmetric recognition of certain gains and losses, accounting for R&D and advertising expenditures, and the absence of information about customers and suppliers in current financial statements imply that the consequences of certain current managerial activities are not reflected in accounting information. This is bias. Similarly, current Gener-

The authors thank two anonymous referees, Gerry Feltham, Thomas Hemmer, Richard Lambert, and workshop participants at the University of British Columbia, University of Chicago, University of Minnesota, University of Washington, and College of William and Mary. The financial support of the Institute of Professional Accounting at the University of Chicago is gratefully acknowledged.

Submitted October 1991.
Accepted January 1993.

ally Accepted Accounting Standards require numerous estimates generating what may be thought of as noise. For example, subjective assessments such as estimated useful lives of assets, loss contingencies, impairment of asset values, and the allocation of purchase price to individual assets and liabilities in corporate acquisitions are likely to introduce noise into accounting information, even in the absence of bias.

A crucial assumption underlying our analysis is that managers allocate effort across many managerial activities, all of which contribute to improving the value of the firm. For example, we can imagine an executive dividing his energies across activities like controlling costs, implementing total quality management initiatives, developing a well-trained workforce, and creating an operating environment where innovation can flourish. While activities such as these have cash flow implications for a firm, we argue that it is difficult, or impossible, to disaggregate the results of operations into individual performance measures that cleanly isolate the effects of the different activities. As a result, compensation contracts in our analysis are based on aggregate performance measures such as accounting income and share price that do not unambiguously distinguish between individual managerial activities.[1] This does not preclude the use of detailed or disaggregated accounting information other than net income as a measure of managerial performance. Rather, it captures the idea that accounting systems cannot realistically measure the economic consequences of all managerial activities in detail. Some aggregation is inevitable, and this imposes a contracting constraint.[2] In our model, limiting the available performance measures to accounting information and share price renders efficient managerial incentives unattainable and creates a potential contracting value for bias and noise.[3]

Given that aggregated accounting information and share price are the only available performance measures, we derive conditions under which *biased* accounting information can be effectively utilized to mitigate the limitations of aggregated measures by better balancing incentives across different managerial activities. In particular, we show that biased accounting information, even if it contains substantial noise, can be better than unbiased accounting information, even if it contains no noise, given that price is also available as a measure of managerial performance. As long as there is a need to provide different incentives for different managerial activities, there is a need for biased accounting information since bias enables the de facto observation of the individual components of output.

---

[1] It is well established that executive compensation contracts are explicitly or implicitly based on both accounting and stock price based measures of managerial performance (e.g., Antle and Smith 1986; Jensen and Murphy 1990). For a discussion of the relative merits of accounting and market measures of performance, see Kaplan and Atkinson (1989) and Watts and Zimmerman (1986).

[2] Indeed, the aggregate nature of most accounting information is often cited as an important constraint in determining the effectiveness of managerial accounting and control systems (see, e.g., Johnson and Kaplan 1987).

[3] The basic idea that the equilibrium price of a firm is an inefficient incentive device in a multi-action setting was first introduced by Paul (1991).

In addition to examining the role of biased accounting information, our analysis demonstrates the benefits of *noisy* accounting information. This suggestion seems counterintuitive since noise imposes risk on the manager without generating any benefits. While the usual intuition holds in most settings, it fails to recognize that there may be an equilibrium relation between accounting information and an endogenously determined share price. A reduction in the level of noise in accounting information motivates investors to curtail private information acquisition, which dilutes the information content of price and thus lessens the usefulness of price as a measure of managerial performance. Therefore, the optimal level of noise in accounting information is a tradeoff between the usefulness of price relative to the usefulness of accounting information as measures of managerial performance.

It is important to emphasize that, in this paper, the benefit of distorted accounting information (bias or noise) that accrues to shareholders is independent of any managerial motivations to manipulate the disclosure system that characterize the "earnings management" literature (see Schipper 1989). Moreover, by focusing exclusively on the stewardship value of information, we show that bias and noise are desirable because they create a "bigger pie" to be shared by all. However, we do not consider the impact of bias and noise on other aspects of shareholder welfare. Bias and noise in accounting disclosures will also impact investors' risk-sharing opportunities both through the direct effect of the distorted public disclosure and the indirect effect on private information acquisition (see, e.g., Diamond 1985). Any overall equilibrium, of course, would have to consider the interaction of all these forces. The exact nature of these tradeoffs remains an unresolved issue.

Key Words: *Stewardship, Managerial compensation, Bias, Noise.*

THE remainder of the paper is organized as follows. Section I describes the basic model and provides a definition of the equilibrium. Section II characterizes the optimal compensation contract and the optimal characteristics of accounting information (bias and noise) in a setting where the market price of the firm is exogenously specified. Section III considers the impact of an endogenous equilibrium market price on the optimal choice of compensation contract and accounting information system. Section IV provides a brief summary.

I. Description of Problem

In this section, we formally describe the model and define its equilibrium. Consider an economy composed of a single risky asset (the firm) and many risk-averse investors, each with negative exponential utility and risk aversion parameter, R.[4] When the economy begins, the firm is established and each investor is endowed with shares of the firm and risk-free bonds. Subsequently, *four* stages (within a single period model) describe the events in this economy.

---

[4] These assumptions allow us to derive a closed form expression for the equilibrium price of the firm in a noisy rational expectations competitive economy.

First, the principal, characterized as a risk-neutral Board of Directors, is given responsibility for designing the manager's wage contract and choosing certain characteristics of the accounting information system.[5] Our intent is not to claim that this formulation accurately captures the actual corporate governance process of the firm, including the Board's influence over the choice of financial reporting systems. Delegating the complete design of the accounting information system to the Board clearly simplifies an extremely complex process. However, whether bias and measurement error actually manifest through the accounting system choices of the Board or through exogenously specified reporting requirements under existing securities regulation is not the issue with which we are specifically concerned. Instead, the primary objective is to utilize the simplification of a risk-neutral Board to demonstrate the existence of a contracting value to bias and measurement error in financial reports and to understand the sources of this value.

Second, the manager, who is precluded from trading in the shares of the firm, takes actions that affect the mean but not the variance of the firm's output.[6] A crucial feature of the model is that managers allocate their effort among multiple activities, necessitating that incentives be provided for all these different activities. Without loss of generality, we limit the analysis to two activities or actions and label them e_1 and e_2.[7] We assume that the firm's output is linear in the two action choices and is given by $\tilde{x}=\tilde{x}_1+\tilde{x}_2$, where $\tilde{x}_1=f\bullet e_1+\tilde{\theta}_1$, $\tilde{x}_2=g\bullet e_2+\tilde{\theta}_2$, and f and g are marginal products of effort. That is:

$$\tilde{x}=(f\bullet e_1+\tilde{\theta}_1)+(g\bullet e_2+\tilde{\theta}_2). \tag{1}$$

$\tilde{x}_1$ and $\tilde{x}_2$ represent the payoffs associated with activities e_1 and e_2. Total mean output generated by the manager is given by $f\bullet e_1+g\bullet e_2$, whereas the total uncertainty in the firm's ouput beyond the control of the manager is given by $\tilde{\theta}_1+\tilde{\theta}_2$. $\tilde{\theta}_1$ and $\tilde{\theta}_2$ are independent normally distributed random variables with zero means and variances of σ_1 and σ_2, respectively. The manager chooses e_1 and e_2 to maximize his utility given the contract and an equilibrium price.

The wage contract is assumed linear in two performance measures: the firm's market price, p, and accounting information, y.

$$w=\alpha+\beta p+\gamma y. \tag{2}$$

The assumption that compensation does not depend explicitly on firm output proxies for the dynamic nature of the problem in the real world, where managers are compensated in the present although output resulting from their efforts may not be fully realized until later periods. The use of a linear contract is supported by the analysis of Holmstrom and Milgrom (1987). They demonstrate that a linear contract is optimal in a dynamic setting that can be equivalently transformed into the static model that we use. Practically, linear contracts provide tractability and are at least consistent with the linear specifcation used in most empirical compensation studies.

---

[5] To be consistent with the risk aversion of shareholders, we could model the Board as risk averse. This would, however, add complexity without adding additional insight. A risk-averse Board would still demand bias and measurement error, although the characterization of its choices would reflect its risk profile.

[6] Allowing the manager to trade the firm's shares influences the incentives that can be provided with accounting and market price based measures of performance. For an analysis of how managerial trading influences these incentives, see Baiman and Verrecchia (1993).

[7] The model could be extended to include many managerial activities and multiple accounting signals as long as the output consequences of the individual managerial activities cannot be disaggregated. For an analysis of multitask principal agent problems, see Holmstrom and Milgrom (1991).

The manager of the firm is risk averse with negative exponential utility and risk aversion parameter r. The opportunity cost of mangerial action is given by $C(e_1,e_2) = C_{11}\cdot(e_1)^2/2+C_{22}\cdot(e_2)^2/2$, implying $C_{12}=0$.[8] The assumptions of linear production and quadratic cost functions facilitate the derivation of closed form expressions for the weights attached to price and accounting information in the compensation contract. In general, a sufficient assumption is that the manager's objective function (described in expression (7) below) is concave in action choices.

Third, each investor j acquires a signal of the form:

$$\tilde{v}_j=\tilde{x}_1+\tilde{x}_2+\tilde{\epsilon}_j, \tag{3}$$

which provides imperfect private information about the total output of the firm. The idiosyncratic errors ϵ_j are independent and identically distributed across traders with mean zero and variance σ_ϵ.[9] The cost of acquiring private information is borne privately by each investor and is given by $k\sigma_\epsilon^{-1}$, with constant marginal cost of k.[10]

At the same time that investors observe their private information, signal y, which we interpret as accounting information, is publicly released by the firm. Accounting information has the form,

$$\tilde{y}=\tilde{x}_1+q\cdot\tilde{x}_2+\tilde{\delta}, \tag{4}$$

and plays two distinct roles in this economy: it is used by investors for valuation purposes and by the Board as a managerial performance measure. While the focus here is on the characteristics of an optimal accounting information system for managerial incentive purposes, the valuation and contracting roles of accounting information are not completely independent. As we will demonstrate in section III below, tailoring the characteristics of accounting information to serve its stewardship role will also affect its valuation role, which may in turn alter the contracting value of equilibrium price through the endogenous responses of traders.

To investigate the existence of a contracting demand for bias and measurement error, we model the Board as potentially choosing two characteristics of accounting information. The first characteristic is the parameter q, which captures *bias* in accounting information; if $q\neq 1$ ($q=1$) in equation (4), accounting information is a biased (unbiased) measure of firm output. For example, when $q=0$, the bias is such that accounting information is completely uninformative about the second component of output $\tilde{x}_2$, and therefore it transmits no information about managerial action e_2. The second characteristic is the parameter σ_δ, the variance of the error term $\tilde{\delta}$, which proxies for various imperfections in accounting information systems. $\tilde{\delta}$ is represented as a normally distributed random variable independent of $\tilde{x}_1$ and $\tilde{x}_2$ with mean zero and variance σ_δ.

---

[8] The assumption that $C_{12}=0$ is made to simplify the analysis. More generally, one can solve this model with a nonzero value for C_{12} without qualitatively changing the results as long as the cost function is convex (i.e., $C_{11}C_{22}>C_{12}^2$).

[9] By pooling their information (the v_js), investors could eliminate their idiosyncratic errors (the ϵ_js) to reveal the total output of the firm (x). However, in a competitive economy, it is reasonable to asssume that equilibrium price is the primary mechanism through which diverse information is aggregated.

[10] In general, a convex function of the form $k(\sigma_\epsilon^{-1})$ is sufficient to establish that investors' choices of σ_ϵ^{-1} will be a maximum

Fourth, stockholders make optimal portfolio decisions by exchanging shares of the firm and risk-free bonds on the basis of private and public information. This exchange of assets implies a price p for the firm determined in a noisy rational expectations equilibrium. Since investors' private information is aggregated in price, price becomes a potentially useful contracting variable. To prevent price from fully revealing investors' private information, we assume that the price is affected by non-output-related events. This is captured by the random variable $\tilde{z}$ (modeled as the supply of the risky asset), where $\tilde{z}$ is independent of all other random events and has mean zero and variance σ_z. The manager is compensated as soon as the market closes.

Note that although these four stages unfold within a single period, the single period nature of the problem is not the crucial feature of the model. Rather, it is the fact that the individual components of output are not observable. This assumption provides contracting value for both accounting disclosures and stock price as imperfect measures of managerial performance. More important, it implies that the compensation contract cannot be based on the individual components of the firm's output $\tilde{x}_1$ and $\tilde{x}_2$. If the individual components of output were observable and thus available for contracting, the individual contract weights placed on $\tilde{x}_1$ and $\tilde{x}_2$ would represent the optimal incentive weights for e_1 and e_2, respectively, unconstrained by aggregation problems. These individual weights, referred to as the "ideal" incentives, are given by:

$$\begin{aligned} &\text{Ideal incentives for } e_1 = \frac{f^2/C_{11}}{f^2/C_{11}+r\sigma_1} \text{ and} \\ &\text{Ideal incentives for } e_2 = \frac{g^2/C_{22}}{g^2/C_{22}+r\sigma_2}. \end{aligned} \tag{5}$$

Expression (5) shows that the two components of output would not, in general, receive equal weights because of differences in the marginal product and marginal cost of the two activities as well as differences in ouput uncertainty. This implies that compensating based on the sum of components $\tilde{x}_1+\tilde{x}_2$, or on the sum of components plus noise such as accounting information or share price, will not in general permit the achievement of ideal incentive weights in expression (5). This inability to achieve ideal incentives represents a loss to the principal that derives from the aggregation process. The analysis in this paper illustrates how particular characteristics of the accounting information system can, in conjunction with share price, alleviate the limitations of writing compensation contracts on aggregated performance measures.

The equilibrium for the model described above is the solution to the Board's problem of selecting an optimal linear compensation contract and accounting system (q or σ_δ or both) subject to: (1) the manager's optimal choice of actions, given the wage contract and the functional form of the equilibrium price; (2) the manager's minimum utility constraint, assumed without loss of generality to be minus one; and (3) a rational market price for the firm determined in a setting where, taking the equilibrium wage contract as given, investors' portfolio choices are optimal and asset markets clear. Formally we have:

$$\max_{\alpha,\beta,\gamma,e_1,e_2,\psi} E_{\tilde{z},\tilde{\delta},\tilde{\theta}_1,\tilde{\theta}_2}[\tilde{x}-(\alpha+\beta\tilde{p}+\gamma\tilde{y})], \tag{6}$$

subject to:

opportunity compensation: $E_{\tilde{z},\tilde{\delta},\tilde{\theta}_1,\tilde{\theta}_2}\left(-\exp\{-r[\alpha+\beta\tilde{p}+\gamma\tilde{y}-C(e_1,e_2)]\}\right)\geq -1,$

rational action choice: $(e_1,e_2)\in \arg\max E_{\tilde{z},\tilde{\delta},\tilde{\theta}_1,\tilde{\theta}_2}\left(-\exp\{-r[\alpha+\beta\tilde{p}+\gamma\tilde{y}-C(e_1,e_2)]\}\right)$, and

rational market price: $\tilde{p}=a_0+a_x\tilde{x}+a_y\tilde{y}-a_z\tilde{z}.$ (7)

ψ in expression (6) represents the choice of accounting information system, where ψ equals either q or σ_δ, or the pair (q,σ_δ). We analyze the Board's problem with a series of results that demonstrate a demand for bias and noise in accounting information. In section II, we analyze the choice of bias and noise when market price is assumed to be exogenously specified. In section III, we endogenize equilibrium market price and examine how this endogeneity can affect the characteristics of an optimal accounting information system.

II. Characteristics of an Optimal Accounting Information System With Exogenous Market Price

This section demonstrates that *biased* accounting information can dominate unbiased accounting information in providing managerial incentives. The analysis ignores the derivation of endogenous market price by assuming that price is exogenously given by:

$$\tilde{p}=\tilde{x}_1+\tilde{x}_2+\tilde{z}. \tag{8}$$

Accounting information continues to be defined by expression (4). The first result characterizes the optimal compensation contract when bias and measurement error are exogenously specified as inherent characteristics of the accounting system. This bias and noise could be thought of as arising from an application of accounting principles promulgated by the SEC and FASB (e.g., accounting for R&D, marketable securities, post retirement benefits, etc.).

Lemma 1: For exogenously specified levels of q and σ_δ, the compensation contract is characterized by the weights on price and accounting information, β and γ:

$$\beta=\frac{r\sigma_\delta\left(\frac{f^2}{C_{11}}+\frac{g^2}{C_{22}}\right)+(q-1)^2\lambda_2\frac{f^2}{C_{11}}+r(q-1)\left(\sigma_2\frac{f^2}{C_{11}}-\sigma_1\frac{g^2}{C_{22}}\right)}{r(\sigma_\delta+\sigma_z)\lambda_1+r(\sigma_\delta+q^2\sigma_z)\lambda_2+(q-1)^2\lambda_1\lambda_2+r^2\sigma_\delta\sigma_z} \text{ and}$$

$$\gamma=\frac{r\sigma_z\left(\frac{f^2}{C_{11}}+q\frac{g^2}{C_{22}}\right)-r(q-1)\left(\sigma_2\frac{f^2}{C_{11}}-\sigma_1\frac{g^2}{C_{22}}\right)}{r(\sigma_\delta+\sigma_z)\lambda_1+r(\sigma_\delta+q^2\sigma_z)\lambda_2+(q-1)^2\lambda_1\lambda_2+r^2\sigma_\delta\sigma_z}, \tag{9}$$

where:

$$\lambda_1 = \frac{f^2}{C_{11}} + r\sigma_1 \text{ and}$$

$$\lambda_2 = \frac{g^2}{C_{22}} + r\sigma_2.$$

Proof: See Appendix.

To clearly illustrate the role played by biased accounting information, consider a limiting case where price is a perfect measure of firm output (i.e., $\sigma_z \rightarrow 0$). This implies that $p = x_1 + x_2$ and, since there is no non-output-related noise remaining in price, it eliminates any role for accounting information in filtering non-output-related noise from price. Setting $\sigma_z = 0$, Lemma 1 shows that despite the absence of non-output-related noise in price, accounting information, even if it contains significant measurement error, will be utilized in the wage contract ($\gamma \neq 0$) unless there is no bias ($q = 1$) or $\sigma_1 g^2/C_{22} = \sigma_2 f^2/C_{11}$.

If $\sigma_z = 0$ and $q = 1$, accounting numbers become redundant and thus useless for contracting purposes given that stock price (which equals $\tilde{x}_1 + \tilde{x}_2$) is already available. Similarly, if $\sigma_z = 0$ and $\sigma_1 g^2/C_{22} = \sigma_2 f^2/C_{11}$, then accounting information again is useless for contracting *regardless* of its bias or measurement error. This follows because the "ideal" incentive weights given in expression (5) are equal in this special case, and, consequently, price, by weighting the two components equally, provides the same incentives as those obtainable when both $\tilde{x}_1$ and $\tilde{x}_2$ are observable separately.

The main point is that as long as the Board desires to differentially weight $\tilde{x}_1$ and $\tilde{x}_2$ in the contract, bias is desirable since it allows the Board to better balance incentives across the two managerial activities. In this extreme case where $p = x_1 + x_2$, bias is the only way to give accounting information contracting value since unbiased accounting information is redundant. Therefore, biased accounting information, even if it possesses substantial measurement error, dominates an unbiased one, even if it possesses no measurement error.

More generally, if $\sigma_z > 0$ and if the Board could control the amount of bias in accounting information, their choice would be characterized by Proposition 1.

Proposition 1: For a given level of σ_δ, the optimal choice of *bias* (q) is given by:

$$q = \frac{\frac{g^2}{C_{22}}\left(\frac{f^2}{C_{11}} + r\sigma_1\right)}{\frac{f^2}{C_{11}}\left(\frac{g^2}{C_{22}} + r\sigma_2\right)} + \frac{r\sigma_\delta\left[(\sigma_1 + \sigma_z)\frac{g^2}{C_{22}} - \sigma_2\frac{f^2}{C_{11}}\right]}{\sigma_z\frac{f^2}{C_{11}}\left(\frac{g^2}{C_{22}} + r\sigma_2\right)}. \tag{10}$$

Proof: See Appendix.

From inspection of equation (10) it is clear that, in general, optimal $q \neq 1$. That is, biased accounting information is desirable. The expression for optimal q is comprised of two terms. The first term is simply the ratio of the "ideal" incentive weights, given in expression (5), that would be placed on x_1 and x_2 individually if they were separately observable. This term captures the ability of bias to mitigate the contracting limitations of having the individual output components $\tilde{x}_1$ and $\tilde{x}_2$ aggregated in price. It is through this term that bias allows for a better balancing of incentives across managerial activities by moving the contract closer to reflecting the ideal incentive weights.

The second term in the expression for q captures the influence of measurement error (σ_δ) in accounting information on the choice of bias. Note that even in the absence of a need to differentiate between activities e_1 and e_2 (i.e., the first term in equation (10) equals 1), the optimal $q \neq 1$ due to the presence of measurement error. Of course, when $\sigma_\delta = 0$ and $\sigma_1 g^2/C_{22} = \sigma_2 f^2/C_{11}$, a compensation contract based on unbiased accounting information ($\tilde{y} = \tilde{x}_1 + \tilde{x}_2$) is as efficient as one based on both $\tilde{x}_1$ and $\tilde{x}_2$.[11] This discussion leads to the following corollary.

Corollary 1: For any σ_δ, when q is chosen optimally, the incentives for e_1 and e_2 given by $\gamma+\beta$ and $q\gamma+\beta$, respectively, are:

$$\gamma+\beta = \frac{\dfrac{f^2}{C_{11}}}{\dfrac{f^2}{C_{11}} + r\sigma_1 + \dfrac{r\sigma_\delta a_z^2 \sigma_z / a_x^2}{(\sigma_\delta + a_z^2 \sigma_z / a_x^2)}} \quad \text{and}$$

$$q\gamma+\beta = \frac{\dfrac{g^2}{C_{22}}}{\dfrac{g^2}{C_{22}} + r\sigma_2}. \tag{11}$$

Proof: Equation (11) is obtained by substituting the optimal q given in Proposition 1 into the expressions for β and γ given by Lemma 1. Q.E.D.

Corollary 1 shows the manner in which the optimal choice of q alleviates the limitations of using share price and accounting information as aggregated measures of managerial performance. Because q directly affects the amount of $\tilde{x}_2$ reflected in accounting information, its optimal choice, in conjunction with the optimal choice of γ and β, provides the Board enough latitude to approach the managerial incentive problem "as if" $\tilde{x}_2$ were observable directly. Consequently, when q is chosen optimally the incentive weight for e_2 is the "ideal" weight given by equation (5).[12] In contrast, since q does not affect the amount of $\tilde{x}_1$ reflected in accounting information, the incentive weight for e_1 will continue to differ from the ideal weight due to the presence of measurement error in accounting information.

Financial reporting requires many estimates to be made, and measurement error, or white noise, is a likely byproduct of most accounting information systems. A standard intuition from agency literature is that measurement error in contracting variables is undesirable, as it represents pure risk without attendant benefits. This intuition implies that, if the Board could freely choose the measurement error in accounting information, they would choose it to be zero.[13] If share price is exogenously given (as assumed in this section), this intuition is right. When $\sigma_\delta = 0$, the second term in the expression for q in equation (10) is zero, but bias remains desirable as long as there is a need to individually tailor the incentives for e_1 and e_2. In addition, when $\sigma_\delta = 0$, Corollary 1 implies that the Board can optimally achieve "ideal" incentives for both e_1 and e_2. The implications of optimally choosing both bias and measurement error in accounting information is summarized in the following proposition.

---

[11] From Proposition 1, it can be shown that unbiased accounting information is desirable if and only if $(\sigma_\delta+\sigma_z)(\sigma_2 f^2/C_{11} - \sigma_1 g^2/C_{22}) = \sigma_\delta \sigma_z g^2/C_{22}$.

[12] The degree of freedom afforded by the free choice of q allows the incentive problem to be solved "as if" the variable multiplying q is directly observable itself. Thus, for example, if q multiplied only the mean of the second component of output given by $g \cdot e_2$ (instead of $\tilde{x}_2$), e_2 in effect would become observable implying a first-best incentive weight for e_2 of one.

Proposition 2: When both bias (q) and measurement error (σ_δ) are chosen optimally, market price is not used in the compensation contract, the incentives for e_1 and e_2 are the ideal second-best incentives, $\sigma_\delta = 0$, and q is given by:

$$q = \frac{\text{Ideal Incentives for } e_2}{\text{Ideal Incentives for } e_1} = \frac{\left(\frac{g^2}{C_{22}}\right) \Big/ \left(\frac{g^2}{C_{22}} + r\sigma_2\right)}{\left(\frac{f^2}{C_{11}}\right) \Big/ \left(\frac{f^2}{C_{11}} + r\sigma_1\right)}. \tag{12}$$

Proof: See Appendix.

Proposition 2 demonstrates that, in general, the Board would choose to eliminate all measurement error but would retain some level of bias in order to balance incentives across activities. In this situation, price is no longer useful as a measure of performance because the contracting problem based on accounting information alone is equivalent to a setting where both $\tilde{x}_1$ and $\tilde{x}_2$ are observable directly. This implies that γ and $q \cdot \gamma$ equal the ideal incentive weights (given by expression [5]) for e_1 and e_2, respectively.

When market price is exogenously specified, its usefulness as a measure of managerial performance is not influenced by investors' reactions to an announcement of accounting information. In the next section, we reexamine the contracting properties of bias and measurement error in a setting where stock price is determined endogenously in equilibrium.

III. Optimal Characteristics of Accounting Information With Endogenous Price

An equilibrium market price adds two new features to the analysis. First, an equilibium price impounds all available value-relevant information in the marketplace, and thus it influences the incremental value of accounting information as an explicit measure of managerial performance. Second, an equilibrium price aggregates investors' private information, which implies that its usefulness as a measure of managerial performance is affected by investors' information acquisition activities. The purpose of this section is to examine how the interaction between an equilibrium price and accounting information can influence the contracting properties of bias and measurement error. In particular, we demonstrate that measurement error in accounting information can have contracting value.

---

[13] While the elimination of all measurement error in accounting estimates is not a realistic possibility, there is likely some degree of control over its realized level. For example, more extensive appraisals, actuarial analyses, or engineering studies could be commissioned to reduce errors in accounting estimates. The objective here is to examine the contracting properties of measurement error as a choice variable in an idealized setting.

We begin by characterizing an equilibrium market price of the form given in equation (7). In a rational expectations equilibrium, investors make conjectures about the relation between prices and information, and the conjectures are self-fulfilling. Let the conjectured relationship between price and investors' information be given by the last constraint in equation (7). This leads to Lemma 2.

Lemma 2: Let $(\hat{e}_1, \hat{e}_2)$ be the manager's equilibrium action choices. When prices are of the form given in the market price constraint in equation (7), a_0, a_x, a_y, and a_z are given by :

$$a_0(1+\beta)=f(\hat{e}_1)+g(\hat{e}_2)-a_x(1+\beta)E(x)-[a_y(1+\beta)+\gamma]E(y)-\alpha,$$

$$a_x(1+\beta)=\frac{\sigma_\epsilon^{-1}+a_x^2/a_z^2\sigma_z}{\mathrm{Var}(x|v_j,y,p)^{-1}},$$

$$a_y(1+\beta)=\frac{\mathrm{Var}(y|x)^{-1}[\mathrm{Cov}(x,y)/\mathrm{Var}(x)]}{\mathrm{Var}(x|v_j,y,p)^{-1}}-\gamma,$$

$$a_z(1+\beta)=\frac{R\sigma_\epsilon(\sigma_\epsilon^{-1}+a_x^2/a_z^2\sigma_z)}{\mathrm{Var}(x|v_j,y,p)^{-1}}, \text{ and}$$

$$\mathrm{Var}(x\,|\,v_j,y,p)^{-1}=\sigma_\epsilon^{-1}+a_x^2/a_z^2\sigma_z+\frac{\sigma_1+q^2\sigma_2+\sigma_\delta}{\sigma_1\sigma_2(1-q)^2+\sigma_\delta(\sigma_1+\sigma_2)}. \tag{13}$$

Proof: The proof follows directly from standard derivations of rational expectations equilibria. See for example, Admati (1985).

The presence of α, β, and γ in the price coefficients in equation (13) reflect the fact that the final value of the firm and thus share price are net of managerial compensation $(\alpha+\beta p+\gamma y)$ paid to the manager. Also, note that the equilibrium price is affected by both bias (q) and measurement error σ_δ. Given this form for share price, we derive the optimal linear contract for any given level of q and σ_δ.

Lemma 3: For any level of q and σ_δ, the weights on price and accounting information in the compensation contract are given by:

$$\beta a_x=\frac{r\sigma_\delta\left(\frac{f^2}{C_{11}}+\frac{g^2}{C_{22}}\right)+(q-1)^2\lambda_2\frac{f^2}{C_{11}}+r(q-1)\left(\sigma_2\frac{f^2}{C_{11}}-\sigma_1\frac{g^2}{C_{22}}\right)}{r(\sigma_\delta+a_z^2\sigma_z/a_x^2)\lambda_1+r(\sigma_\delta+q^2a_z^2\sigma_z/a_x^2)\lambda_2+(q-1)^2\lambda_1\lambda_2+r^2\sigma_\delta a_z^2\sigma_z/a_x^2}, \text{ and}$$

$$\gamma+\beta a_y=\frac{ra_z^2\sigma_z/a_x^2\left(\frac{f^2}{C_{11}}+q\frac{g^2}{C_{22}}\right)-r(q-1)\left(\sigma_2\frac{f^2}{C_{11}}-\sigma_1\frac{g^2}{C_{22}}\right)}{r(\sigma_\delta+a_z^2\sigma_z/a_x^2)\lambda_1+r(\sigma_\delta+q^2a_z^2\sigma_z/a_x^2)\lambda_2+(q-1)^2\lambda_1\lambda_2+r^2\sigma_\delta a_z^2\sigma_z/a_x^2} \tag{14}$$

where:

$$\lambda_1=\frac{f^2}{C_{11}}+r\sigma_1 \text{ and}$$

$$\lambda_2=\frac{g^2}{C_{22}}+r\sigma_2.$$

Proof: See Appendix.

The weights on accounting information and price in Lemma 3 parallel those of Lemma 1 where price was exogenously specified, with two important differences. First, the total incentive weight on accounting information is the sum of the explicit weight γ and an implicit weight βa_y that is created by the impounding of accounting information in equilibrium market price. Second, the noise in price in Lemma 3 is captured by the term $a_z^2\sigma_z/a_x^2$, where a_z/a_x is endogenously determined by investors' information acquisition decisions.[14] The parallel between Lemmas 1 and 3 can be seen clearly from a simple transformation of equilibrium price. The transformation based on common knowledge, $(p-a_0-a_y y)/a_x$, yields the transformed price variable $x_1+x_2+(a_z/a_x)z$. This is identical in form to the exogenous price variable specified in equation (8) above, with the exception that the noise term in equation (8) is exogenous and here it is endogenous through the term a_z/a_x.

The endogeneity of noise in price implies that choices made about the informational properties of accounting information can also affect the informational properties of price. This follows because the characteristics of accounting information directly influence investors' private information acquisition decisions. Therefore, optimal choices of bias and noise must consider the impact on investors' private information acquisition decisions and the resulting noise in price. Of course, to the extent that price is useless for contracting, its endogeneity becomes irrelevant. For example, if the Board can choose both bias and measurement error, they would choose σ_δ and q exactly as in Proposition 2 of section III in order to achieve ideal incentive weights. Thus in this case, even though σ_δ and q affect private information acquisition, private information and the capital market have no bearing on the agency problem.

The preceding discussion indicates that if both q and σ_δ are choice variables, σ_δ will always be chosen equal to zero. Therefore, to investigate further the contracting properties of noise, we make the assumption that the bias (q) in accounting information is exogenously fixed.[15] This limits the Board's discretion in designing the accounting information system to the choice of measurement error (σ_δ). Given this setting, we demonstrate that there are situations where $\sigma_\delta=0$ is not optimal, given that market price is endogenous.

The analysis exploits the endogenous relation between the information content of accounting information, as captured by σ_δ, and the private information acquisition decisions of traders. The value of share price as a measure of managerial performance depends directly on these private information acquisition decisions, and these decisions depend on the level of noise, σ_δ, in earnings. This follows from the fact that share price reflects investors' private information about firm value and that investors consider all publicly available information, such as accounting disclosures, when deciding on how much costly private information to acquire.

Private information acquisition is characterized as the observation of a signal $v_j=x_1+x_2+\epsilon_j$ with precision (conditional on x) of σ_ϵ^{-1} at a cost of $k\bullet\sigma_\epsilon^{-1}$. Since in our model public and private information are substitutes, reducing measurement error in accounting information reduces the precision of the private information acquired by investors, which in turn reduces the informativeness of price. Thus, the choice of σ_δ must trade off the increased contracting value of a less noisy accounting measure against the reduced information content of price.

---

[14] In contrast, in section II, the noise in price is exogenously specified to be σ_z.

[15] Alternatively, we could assume that there are some constraints on the level of q that can be chosen. These constraints may arise from GAAP itself, where mandated measurement rules can impose an inescapable level of bias.

As discussed above, equilibrium price can be transformed into a random variable with form $x_1+x_2+(a_z/a_x)z$. This implies that the ability of price to transmit information about payoff is determined by the size of a_z/a_x. That is, the extent to which investors' private information is reflected in price can be captured by the term $a_x^2/a_z^2\sigma_z$, which we refer to as the informativeness of price. This term is a function of the precision of private signals acquired, σ_ϵ^{-1}, which in turn is a function of σ_δ. We summarize the preceding discussion in the following lemma and show that the informativeness of price is a nondecreasing function of σ_δ.

Lemma 4: The informativeness of price, $a_x^2/a_z^2\sigma_z$, is a nondecreasing function of the measurement error in accounting information.

Proof: See Appendix.

Lemma 4 implies that reducing measurement error in accounting information can exacerbate the agency problem by reducing the usefulness of share price as a measure of managerial performance. The implication is that the optimal choice of σ_δ represents a tradeoff between the contracting benefit of less measurement error in accounting information and the contracting sacrifice of a less informative price, as long as share price remains a useful measure of managerial performance.

As might be expected, precisely characterizing the optimal choice of σ_δ is a complicated task. Situations under which a non-zero σ_δ is optimal are conditioned on complex relationships among the exogenous parameters of the model. However, while the precise conditions under which $\sigma_\delta>0$ would be chosen are difficult to characterize, it is possible to generate parametric examples that clearly reveal the essence of the economic tradeoffs involved. Consider first a situation where there is no exogenous bias in accounting information (i.e., $q=1$). In this setting, optimal $\sigma_\delta=0$ since $y=\tilde{x}_1+\tilde{x}_2$ completely dominates share price as a contracting variable and thus renders the amount of information reflected in price irrelevant to the contracting problem. In contrast, when exogenous $q\neq 1$, there exists levels of the exogenous parameters such that $\sigma_\delta=0$ is not optimal.

We illustrate the economic forces underlying the desirability of measurement error in accounting information with a series of three related examples. In all three examples, exogenous $q=0$ so that accounting information is completely uninformative about the second component of the firm's output x_2, and in addition, $f=1$, $\sigma_z=2$, $k=0.4$, and all other parameters in the model are equal to one.[16] The three examples examine the optimal choice of σ_δ for three different values of g, the marginal product of e_2. Because q is set equal to zero exogenously, share price is the only means of providing incentives for the second managerial activity, e_2.[17] By manipulating the level of marginal product of e_2, g, relative to the marginal product of e_1, f, we determine the relative importance of motivating e_2 in the wage contract. As the examples show, the more important motivating e_2 becomes, the more the Board desires to retain information content in price, and so the greater the level of σ_δ chosen.

The solutions to the three examples are shown in graphical form in figure 1. The graphs plot the net profit of the firm after managerial compensation as a function of the

---

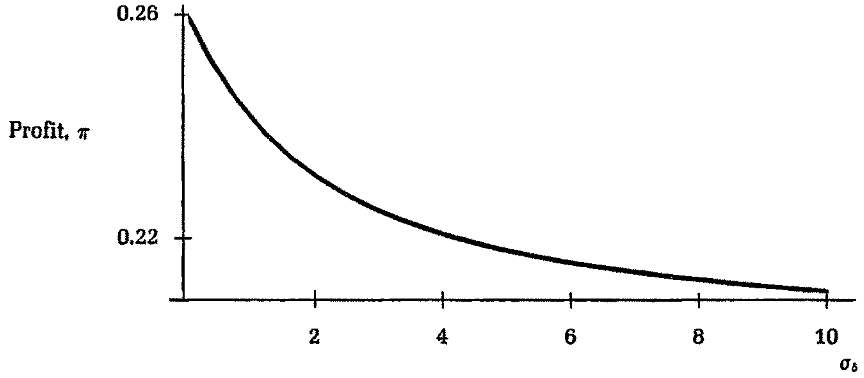
[16] f represents the marginal product of e_1, σ_z represents the uncertainty in the supply of the risky asset that prevents price from being fully revealing, and k represents the marginal cost of a unit of precision of private information.

[17] The contracting implications of accounting information where $q=0$ are extensively analyzed in Bushman and Indjejikian (1993).

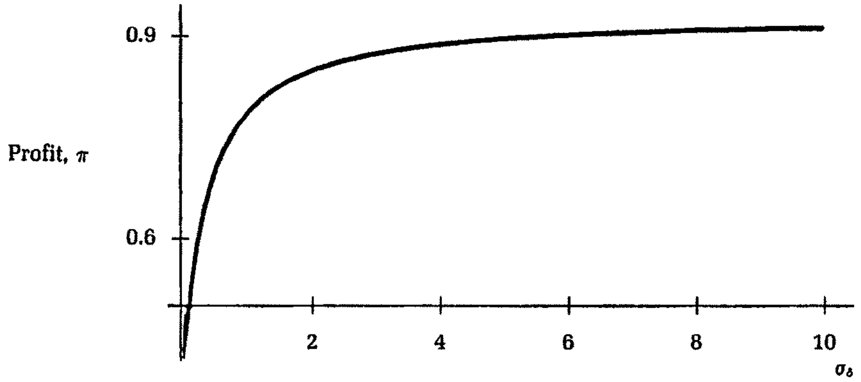
Figure 1

Parameter Values:
$r=R=C_{11}=C_{22}=\sigma_1=\sigma_2=1$; $k=0.4$; $\sigma_\varepsilon=2$; $q=0$

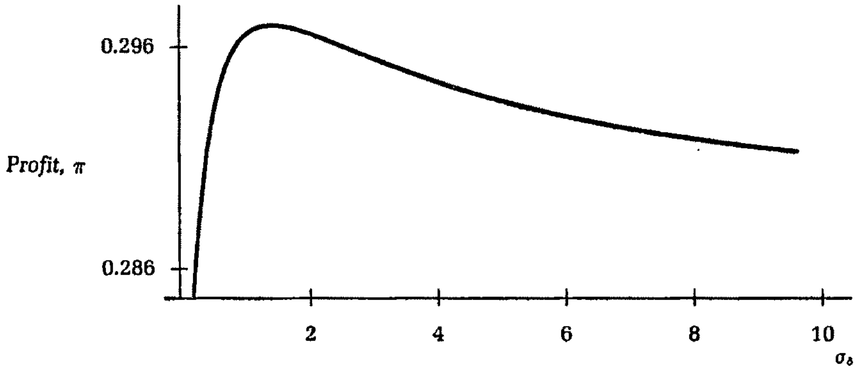
Panel A. $f=1$; $g=1$



Panel B. $f=1$; $g=2$



Panel C. $f=1$; $g=1.2$



choice parameter σ_δ.[18] Consider first panel A of figure 1 in which $f=g=1$. In this case, the marginal products of e_1 and e_2 are identical resulting in optimal $\sigma_\delta=0$. In essence, because the marginal products are equal, the Board is willing to sacrifice some ability to motivate e_2 by increasing noise in price in order to perfectly observe x_1. This allows them to achieve ideal incentives for e_1.

Next, consider panel B, where $f=1$ and $g=2$ implies that e_2 incentives are far more important then e_1. In this case, the optimal solution is to set $\sigma_\delta=\infty$. The higher e_2 marginal product (high g) increases the contracting value of price as the sole means of generating e_2 incentives for the manager, As a result, σ_δ is chosen to destroy the information content of earnings (i.e., $\sigma_\delta=\infty$) which maximizes the information content of price by motivating maximum private information acquisition by informed traders.

Last, panel C reflects an intermediate case with $f=1$ and $g=1.2$. This situation results in $0<\sigma_\delta<\infty$, and as a result the Board strikes a compromise between e_1 and e_2 incentives. By adding a moderate level of noise, the Board maintains an ability to provide e_2 incentives without completely destroying e_1 information in the accounting report. The choice of $\sigma_\delta>0$ is dictated by the contracting value of increasing the information content of price relative to increasing the information content of accounting reports.

The idea that an accounting information signal can dilute the usefulness of share price as a measure of managerial performance is closely related to the results in Dye (1985), section IV. Dye (1985) shows that when information is released early by a firm, the fact that it will be immediately impounded in price preempts future rounds of trading and so prevents the superior contracting information possessed privately by traders from being impounded in price. The basic economic force underlying Dye's result is the inability of traders to commit to refrain from trading based on the release of public information and thereby prevent its reflection in price. In our model, it is investors' inability to refrain from responding to more precise public information by modifying the amount of private information acquired that leads to a dilution of the usefulness of share price, and thus of the value of noise.

IV. Summary

This paper identifies a shareholder demand for "distorted" accounting information systems. Focusing on the stewardship role of accounting information, we show that when a manager of a firm is engaged in more than one activity, the need to provide tailor-made incentives for each activity usually leads to a demand for "distorted" information. This is not to suggest that firms or standard setters engage in deliberate distortion of accounting information. Instead, our results are meant to illustrate the potentially irreconcilable difficulties facing the FASB, or any accounting standard setting body, in setting general purpose accounting standards.

The crucial assumption underlying our results is the idea that managerial compensation contracts are based on one or more measures of aggregate firm value such as accounting information and share price and not on the individual components of a firm's output. The aggregate nature of these types of measures usually prevents the provision of efficient managerial incentives. Therefore, to the extent that shareholders can exercise some discretion over the choice of accounting information systems, they will choose one that is biased and perhaps noisy in the interest of providing efficient managerial incentives.

---

[18] The profit function is derived by substituting the optimal wage contract coefficients from equation (14) back into the Board's objective function, and in addition endogenizing a_t/a_x by substituting in the traders' optimal choice of σ_ϵ as a function of σ_δ as given in the proof of Lemma 4.

Several economic forces that could affect our results are not considered here. We do not consider managerial motivations to manipulate earnings for private gain in the presence of asymmetric information. We also do not consider how the public release of accounting information affects individual risk-averse shareholders by changing their risk-sharing opportunities. The extent to which these economic forces exacerbate or mitigate the demand for distorted accounting information analyzed in this paper remains an open question.

Appendix

Proof of Lemma 1

An optimal linear compensation design offers the manager a contract that maximizes the expected payoff of the firm less expected wages paid to the manager. The programming problem presented in equations (6) and (7) can be restated by exploiting the tractability of the constraints. Solving the opportunity compensation constraint in equation (7) for α and substituting into the objective function allows expected compensation to be written as the sum of the manager's cost of choosing e_1 and e_2 given by $C(e_1,e_2)$, and the cost of risk borne by the manager given by $(r/2)\mathrm{Var}(w)$. We also substitute the agent's first order conditions for the rational action choice constraint in equation (7), as these conditions describe a unique optimum for given γ, β, and p. Hence, the problem defined by equations (6) and (7) for a given level of q and σ_δ can be rewritten as:

$$\max_{\beta,\gamma,e_1,e_2} \quad [f\cdot e_1+g\cdot e_2-C(e_1,e_2)-(r/2)\mathrm{Var}(\alpha+\beta p+\gamma y)] \tag{15}$$

subject to:

$$(\beta+\gamma)f=C_1(e_1,e_2), \text{ and}$$
$$(\beta+q\gamma)g=C_2(e_1,e_2).$$

The first order conditions are given by:

I. $f\mu_1+g\mu_2-r\beta\mathrm{Var}(p)-r\gamma\mathrm{Cov}(p,y)=0,$

II. $f\mu_1+qg\mu_2-r\gamma\mathrm{Var}(y)-r\beta\mathrm{Cov}(p,y)=0,$

III. $(f-C_{11}e_1)-\mu_1C_{11}=0,$

IV. $(g-C_{22}e_2)-\mu_2C_{22}=0,$

V. $(\beta+\gamma)f-C_{11}e_1=0,$ and

VI. $(\beta+q\gamma)g-C_{22}e_2=0.$ (16)

The simultaneous solution to the above expressions gives β and γ in Lemma 1. The concavity of the program ensures that the solution is a maximum.

Proof of Proposition 1

From the solution to the set of equations (16), the objective function can be rewritten as follows:

$$\begin{aligned}\pi(q,\sigma_\delta)=&(\gamma+\beta)\left(\frac{f^2}{C_{11}}+r\sigma_1\right)\left(\frac{f^2/C_{11}}{f^2/C_{11}+r\sigma_1}-\frac{(\gamma+\beta)}{2}\right)\\&+(q\gamma+\beta)\left(\frac{g^2}{C_{22}}+r\sigma_2\right)\left(\frac{g^2/C_{22}}{g^2/C_2+r\sigma_2}-\frac{(q\gamma+\beta)}{2}\right)\\&-r\sigma_z\frac{\beta^2}{2}-r\sigma_\delta\frac{\gamma^2}{2},\end{aligned} \tag{17}$$

where β and γ are given by Lemma 1. The first order condition with respect to q gives rise to the following two stationary points:

$$q_1=\frac{\frac{g^2}{C_{22}}\left(\frac{f^2}{C_{11}}+r\sigma_1\right)}{\frac{f^2}{C_{11}}\left(\frac{g^2}{C_{22}}+r\sigma_2\right)}+\frac{r\sigma_\delta\left[(\sigma_1+\sigma_x)\frac{g^2}{C_{22}}-\sigma_2\frac{f^2}{C_{11}}\right]}{\sigma_x\frac{f^2}{C_{11}}\left(\frac{g^2}{C_{22}}+r\sigma_2\right)}\ \text{and}$$

$$q_2=\frac{\left(\sigma_1\frac{g^2}{C_{22}}-\sigma_2\frac{f^2}{C_{11}}\right)-\sigma_x\frac{f^2}{C_{11}}}{\left(\sigma_1\frac{g^2}{C_{22}}-\sigma_2\frac{f^2}{C_{11}}\right)+\sigma_x\frac{g^2}{C_{22}}}. \tag{18}$$

By the second order conditions, it can be shown q_1 is a relative maximum whereas q_2 is not a maximum. It can also be shown that the objective function at $q=q_1$ dominates all corner solutions for q. Therefore, it follows that $q=q_1$ is a global maximum.

Proof of Proposition 2

For the model presented in this paper, contracts based on the observability of $\tilde{x}_1$ and $\tilde{x}_2$ (weakly) dominate all second-best contracts based on various aggregate measures of $\tilde{x}_1$ and $\tilde{x}_2$ (such as accounting information and share price). It can be shown that when $\tilde{x}_1$ and $\tilde{x}_2$ are observable, the net profit of the firm (after managerial compensation) is given by:

$$\frac{1}{2}\frac{(f^2/C_{11})^2}{f^2/C_{11}+r\sigma_1}+\frac{1}{2}\frac{(g^2/C_{22})^2}{g^2/C_{22}+r\sigma_2}. \tag{19}$$

The proof of this proposition proceeds by showing that setting q as in equation (12) and $\sigma_\delta=0$, one can construct a compensation contract with β and γ given by Lemma 1 to achieve the second-best optimal profit given by equation (19). When $\sigma_\delta=0$ and q is given by equation (12), Lemma 1 implies that:

$$\beta=0\ \text{and}$$

$$\gamma=\frac{f^2/C_{11}}{f^2/C_{11}+r\sigma_1}. \tag{20}$$

At these values of q, σ_δ, β, and γ, it can be shown that the net profit of the firm given by equation (17) equals the net profit that would be achieved if $\tilde{x}_1$ and $\tilde{x}_2$ were observable directly given by equation (19).

Proof of Lemma 3

The proof of Lemma 3 parallels the proof of Lemma 1 except that the expression for price has the (equilibrium) linear form given by equation (7). For algebraic convenience, we define the contract choice variables as βa_x and $\gamma+\beta a_y$, the actual incentive weights on the private information component of price and accounting income respectively, rather than on β and γ alone. Hence, the programming problem for a given level of q and σ_δ can be rewritten as follows:

$$\max_{\beta a_x,\gamma+\beta a_y,e_1,e_2}\ \{f\cdot e_1+g\cdot e_2-C(e_1,e_2)-(r/2)\mathrm{Var}[\alpha+\beta a_x(p-a_y y)/a_x+(\gamma+\beta a_y)y]\} \tag{21}$$

subject to:

$$[\beta a_x+(\beta a_y+\gamma)]f=C_1(e_1,e_2)\ \text{and}$$

$$[\beta a_x+q(\beta a_y+\gamma)]g=C_2(e_1,e_2).$$

The first order conditions are given by:

I. $f\mu_1+g\mu_2-r\{\beta a_x\mathrm{Var}[(p-a_y y)/a_x]+(\gamma+\beta a_y)\mathrm{Cov}[(p-a_y y)/a_x,y]\}=0,$

II. $f\mu_1+qg\mu_2-r\{(\gamma+\beta a_y)\mathrm{Var}(y)+\beta a_x\mathrm{Cov}[(p-a_y y)/a_x,y]\}=0,$

III. $(f-C_{11}e_1)-\mu_1 C_{11}=0,$

IV. $(g-C_{22}e_2)-\mu_2 C_{22}=0,$

V. $[\beta a_x+(\gamma+\beta a_y)]f-C_{11}e_1=0,$ and

VI. $[\beta a_x+q(\gamma+\beta a_y)]g-C_{22}e_2=0.$ (22)

The simultaneous solution to the above expressions gives βa_x and $(\gamma+\beta a_y)$ in Lemma 3. The concavity of the program ensures that the solution is a maximum.

Proof of Lemma 4

When the cost of acquiring private information is given by $k\sigma_\epsilon^{-1}$ with constant marginal cost of k, the choice of σ_ϵ^{-1} for all investors can be characterized as follows:

$$\sigma_\epsilon^{-1}=\text{Max}\left(0,\ \sigma_\epsilon^{-1}\left|\frac{1}{2kR}=\sigma_\epsilon^{-1}+\frac{a_z^2}{a_x^2\sigma_x}+\frac{\sigma_1+q^2\sigma_2+\sigma_\delta}{\sigma_\delta(\sigma_1+\sigma_2)+\sigma_1\sigma_2(1-q)^2}\right.\right), \tag{23}$$

where $a_{x2}/a_x^2\sigma_x=R^{-2}\sigma_\epsilon^{-2}/\sigma_x$. Expression (23) follows from standard derivations of rational expectations equilibria with endogenous private information acquisition (see, e.g., Verrecchia 1982). Using equation (23) we have:

$$\frac{d(\sigma_\epsilon^{-1})}{d\sigma_\delta}=\frac{\sigma_\epsilon^{-1}}{(\sigma_\epsilon^{-1}+2a_z^2/a_x^2\sigma_x)}\frac{(\sigma_1+\sigma_2 q)^2}{[\sigma_\delta(\sigma_1+\sigma_2)+\sigma_1\sigma_2(q-1)^2]^2}\geq 0, \tag{24}$$

which implies that:

$$\frac{d(a_z^2/a_x^2\sigma_x)}{d\sigma_\delta}=\frac{2R^{-2}\sigma_\epsilon^{-1}}{\sigma_x}\frac{d\sigma_\epsilon^{-1}}{d\sigma_\delta}\geq 0. \tag{25}$$

References

Admati, A. R. 1985. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica* (May): 629–57.

Antle, R., and A. Smith. 1986. An empirical investigation into the relative performance evaluation of corporate executives. *Journal of Accounting Research* (Spring): 1–39.

Baiman, S., and R. E. Verrecchia. 1993. Earnings and price based compensation contracts in the presence of discretionary trading. Working paper, University of Pennsylvania.

Bushman, R., and R. J. Indjejikian. 1993. Accounting income, stock price and managerial compensation. *Journal of Accounting & Economics*. Forthcoming.

Diamond, D. W. 1985. Optimal release of information by firms. *Journal of Finance* (September): 1071–94.

Dye, R. A. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* (Spring): 123–45.

Financial Accounting Standards Board. 1980. Qualitative characteristics of accounting information. *Statement of Financial Accounting Concepts No. 2* (May). Stamford, CT: FASB.

Gjesdal, F. 1981. Accounting for stewardship. *Journal of Accounting Research* (Spring): 208–31.

Holmstrom, B., and P. Milgrom. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* (March): 303–28.

——— and ———. 1991. Multitask principal-agent analyses: Incentive contacts, asset ownership and job design. *Journal of Law, Economics & Organization* 7 (Supplement): s24–s52.

Jensen, M. C., and K. J. Murphy. 1990. Performance pay and top management incentives. *Journal of Political Economy* (April): 225–64.

Johnson, H. T., and R. Kaplan. 1987. *Relevance Lost: The Rise and Fall of Management Accounting*. Harvard Business School Press, Boston, MA.

Kaplan, R., and A. Atkinson. 1989. *Advanced Management Accounting*. Englewood Cliffs, NJ: Prentice Hall.

Paul, J. M. 1991. On the efficiency of stock based compensation. Working paper, University of Michigan.

Schipper, K. 1989. Earnings management. *Accounting Horizons* (December): 91–102.

Verrecchia, R. E. 1982. Information acquisition in a noisy rational expectations economy. *Econometrica* 50, (6): 1415–30.

Watts, R., and J. L. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice Hall.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 783-803

Auditors' Assessments of Inherent and Control Risk in Field Settings

William S. Waller
University of Arizona

SYNOPSIS AND INTRODUCTION: Current policy on how auditors should limit uncertainty about misstatements in auditee assertions is based on the audit risk model (AICPA 1992) that decomposes the components of audit risk as inherent risk (*IR*), control risk (*CR*), and detection risk (*DR*). The literature on the audit risk model has focused on *a priori* analyses of the model's assumptions and implications (see, e.g., Cushing and Loebbecke 1983; Kinney 1983, 1989, 1992; Leslie 1984), and auditors' risk assessments in experimental settings (see, e.g., Colbert 1988; Daniel 1988; Jiambalvo and Waller 1984; Libby et al. 1985). Absent are empirical studies that examine applications of the model in field settings.

This article reports empirical evidence on auditors' *IR* and *CR* assessments in field settings by analyzing archival data drawn from the audit workpapers of KPMG Peat Marwick. As a part of audit planning, the firm requires its auditors to make and document *IR* and *CR* assessments for each assertion of each significant account.[1] The assessments are made with respect to tolerable error, an algorithm-based measure of planning materiality at the assertion level.[2] The data include approximately 5,000 risk

---

[1] Five assertions are pertinent for each account: completeness (all exchanges that should be recorded are recorded), existence (all exchanges recorded by the accounting system occurred and were valid), accuracy (all exchanges recorded by the accounting system were recorded accurately), valuation (exchanges reflected in the financial statements are valued appropriately and in accordance with relevant accounting policies), and ownership (auditee has appropriate rights and obligations as to financial statement items). This set of assertions varies from those stated in the professional standards (AICPA 1992, AU 326) in that the firm requires IR and CR assessments for both accuracy and valuation assertions, instead of just the accuracy assertion of the professional standards. The firm also requires consideration, but not explicit risk assessments, of assertions regarding presentation and disclosure.

[2] The computation of tolerable error involves two steps. To compute planning materiality at the financial statement level (gauge), a formula is applied to the greater of sales or total assets. To compute tolerable error at

This study was conducted while the author was Audit Research Fellow with the Audit Research Group of KPMG Peat Marwick. Special thanks go to John Willingham, Tim Bell, and other Group members. Comments on earlier drafts by two anonymous reviewers, Bill Felix, Bill Kinney, Brian Shapiro, Ira Solomon, and participants of the 1992 AAA Doctoral Consortium and research workshops at Alberta, Arizona, Colorado, Connecticut, Florida, Georgia, Indiana, Iowa, North Texas, and Wisconsin, as well as the research assistance of Audrey Gramling and Mark Zimbelman, are greatly appreciated.

Submitted April 1991.
Accepted March 1993.

assessments at the assertion level for trade accounts receivable, inventory, and trade accounts payable, on a total of 215 audit engagements. The data also indicate, for each assertion, whether a misstatement exceeding tolerable error was detected by the auditor.

The data analysis considers four issues, the first of which is whether there is a statistical association between auditors' *IR* and *CR* assessments. *A priori* researchers (e.g., Cushing and Loebbecke 1983) argue that the audit risk model's multiplicative combination of *IR* and *CR* suggests independence between risk components, which contradicts auditors' conventional wisdom of dependence. The analysis in this study concludes that the dependence problem arises because (1) its event structure is ill-defined and (2) it fails to recognize that an auditor's assessments are conditional on his or her knowledge. A knowledge-based dependence may produce a statistical association between *IR* and *CR*. Contrary to expectation, the empirical evidence supports the conclusion of an insignificant association between *IR* and *CR*; however, in a predominance of cases, *CR* is assessed at the maximum probably for reasons of efficiency.

The remaining issues pertain to the policy requirement of *assertion-level* risk assessments. Viewed generally over many audit engagements, current policy depends on the following premises: (1) the rate of misstatements varies over assertions, (2) auditors' risk assessments vary over assertions, and (3) the association between the rate of misstatements and auditors' risk assessments is positive (i.e., the assessments are accurate). The data analysis examines the empirical validity of each premise, and supports the first inasmuch as there are significant differences in the rate of *detected* misstatements over assertions for each account. However, auditors typically assess *IR* and *CR* at the same value for all assertions for an account, which is inconsistent with the second premise. When the data pertaining to all assertions for an account are included in the analysis, the association between *IR* and the rate of detected misstatements (after controlling for *CR* and *DR*) tends to be positive but low, which indicates modest support for the third premise. When the analysis includes only the "most important" assertion for each account, the association is considerably stronger. In general, auditors' risk assessments are consistent with a heuristic that deliberately assesses risk for an account's "most important" assertion but does so mechanically for other assertions. Taken as a whole, the results indicate a need either to reconsider the policy of multiple risk assessments for an account or to enhance auditors' ability to assess assertion-specific risk.

Key Words: *Audit risk, Inherent risk, Control risk, Field-based research.*

Data Availability: *Contact the author.*

---

the account and assertion levels, another formula is applied that incorporates gauge and the account balance. For example, suppose an auditee has sales of $1,000,000, total assets of $3,000,000, and trade accounts receivable of $250,000. In this case, gauge is $38,300, and tolerable error is $11,000 for each assertion of trade accounts receivable.

THE remainder of the article is organized as follows. I develop the four research issues in section I and state corresponding empirical propositions, relying on a probabilistic view of audit risk consistent with behavioral decision analysis. The data analysis with respect to the four propositions is presented in section II. The results and possible future research are discussed in section III.

I. Research Issues

When applying the audit risk model at the assertion level, the auditor begins by specifying a target value of audit risk (AR), the risk that the assertion contains a misstatement exceeding tolerable error at the end of the audit. To achieve this target, the model decomposes AR into occurrence risk (OR), the risk that the assertion contains a misstatement prior to the audit, and detection risk (DR), the risk that the auditor fails to detect a realized misstatement, where $AR = OR \times DR$. Given the auditor's OR assessment, test procedures are selected such that $DR = AR/OR$. Some firms additionally decompose OR into inherent risk (IR) and control risk (CR), where $OR = IR \times CR$.[3] As defined in professional standards, IR is the risk that the assertion contains a misstatement, assuming there are no related controls, and CR is the risk that controls fail to prevent or detect a misstatement. In this case, the auditor assesses IR and CR and selects test procedures such that $DR = AR/(IR \times CR)$.

Professional standards state that the audit risk model "expresses the general relationship of the risks associated with the auditor's assessment of inherent and control risks" and "is not intended to be a mathematical formula including all factors that may influence the determination of individual risk components" (AICPA 1992, AU 350.48). Thus, it is debatable whether logical exactness should be attributed to the model. In conformance with past research, the view adopted in this study is that policy regarding audit risk should be based on a rigorous foundation, such as probability theory.[4]

In this regard, it is useful to compare the audit risk model with behavioral decision analysis (Raiffa 1968; Von Winterfeldt and Edwards 1986). The latter operationalizes expected utility theory by decomposing an individual's decision problem into the tasks of utility and probability assessment. Regarding probability assessment, decomposition often extends to the explicit incorporation of component events and conditioning information. Decomposition is guided throughout by two principles. First, decomposition relies on a theory-based link between components, which permits a coherent integration of the separate assessments (e.g., the expectation operator for combining utility and probability assessments, or probability rules for combining assessments of component events). Second, the practical value of decomposition depends on the overall gain in effectiveness from performing two or more assessment tasks, each with a relatively narrow focus, versus a holistic approach. Without such gain, a decomposition that merely creates more tasks is costly.

---

[3] Professional standards state that an auditor may make separate or combined assessments of IR and CR (AICPA 1992, AU 312.24). Also, some firms decompose DR into analytical procedures risk (AP), the risk that analytical procedures fail to detect a realized misstatement, and test of details risk (TD), the risk that tests of details fail to detect an otherwise undetected misstatement, where $DR = AP \times TD$.

[4] Professional standards do not clarify the relation between audit risk and probability. The word "probability" does not appear in various sections that are directly relevant to audit risk, including AU 312, "Audit Risk and Materiality in Conducting an Audit"; AU 316, "The Auditor's Responsibility to Detect and Report Errors and Irregularities"; and AU 350, "Audit Sampling." However, in AU 319, "Consideration of the Internal Control Structure in a Financial Statement Audit," the definition of maximum control risk is given as the "greatest probability" that a misstatement will not be prevented or detected by an auditee's controls.

Figure 1
Event Sequence

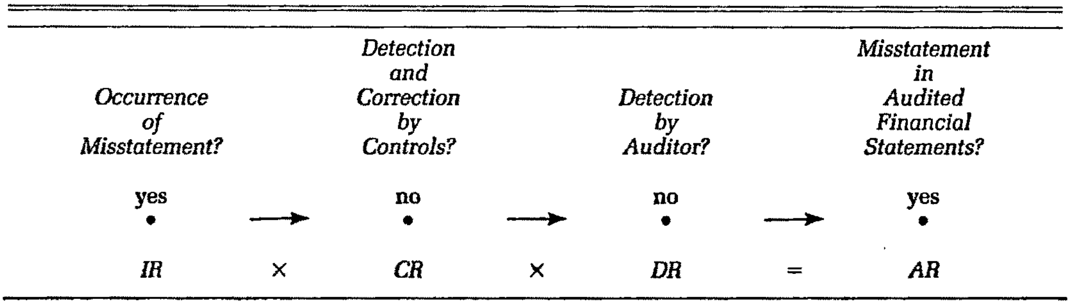
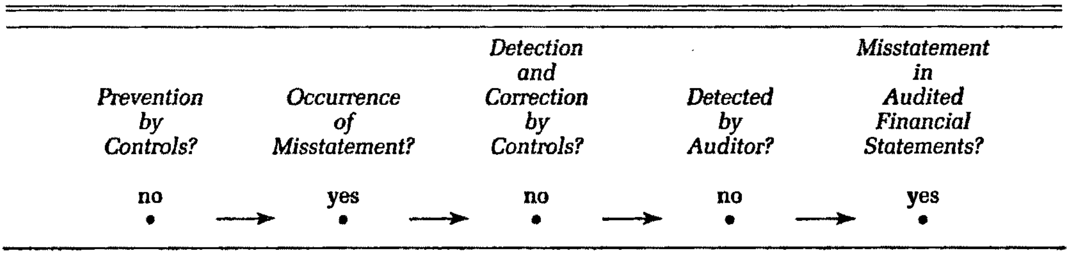

Occurrence of Misstatement?
Detection and Correction by Controls?
Detection by Auditor?
Misstatement in Audited Financial Statements?
yes
no
no
yes
IR
×
CR
×
DR
=
AR

Figure 2
Event Sequence with Prevention by Controls


Prevention by Controls?
Occurrence of Misstatement?
Detection and Correction by Controls?
Detected by Auditor?
Misstatement in Audited Financial Statements?
no
yes
no
no
yes

In contrast with behavioral decision analysis, the audit risk model's decomposition is not theory-based, which raises doubt about whether its components can be integrated as intended (e.g., $IR \times CR$ may not equal OR). Further, the practical value of the model's decomposition is lost when a strong dependence exists between *IR* and *CR* assessments for an assertion, or among assessments over assertions, such that the separate assessments are informationally redundant.

Association Between IR *and* CR *Assessments*

The audit risk model seems to rely on the event sequence in figure 1 (cf. Graham 1985). That is, a misstatement may or may not occur, which relates to *IR*. When a misstatement occurs, controls may or may not detect (and correct) it, which relates to *CR*. When a misstatement is not detected by controls, the auditor may or may not detect it, which relates to *DR*. When a misstatement is not detected by controls or the auditor, there is a misstatement in the audited financial statements, which relates to *AR*. A problem with this sequence is its failure to include the preventive effect of controls; CR depends on both the preventive and detective effects of controls. The event sequence in figure 2 incorporates the preventive effect, but note that *IR* would depend on the quality of controls, contrary to the professional standards' definition of IR. Apparently, there is a fundamental dependence between *IR* and *CR* that is not captured by the audit risk model (Cushing and Loebbecke 1983; Kinney 1984; Leslie 1984).

To show why the dependence problem arises, consider a probabilistic view of audit risk consistent with behavioral decision analysis.[5] When formalized, the event structure in figure 1 has three components: M, the event of a misstatement ($-M$ is its complement); C, the event of detection (and correction) of a misstatement by controls; and D, the event of detection of a misstatement by the auditor. Thus, a misstatement in audited financial statements requires the conjunction of M, $-C$, and $-D$. For a component event, say M, an auditor's subjective probability of M is conditioned by his or her knowledge at the time of assessment, K, and is represented as $P_K(M)$. For two component events, say C and M, $P_K(C \mid M)$ represents an auditor's subjective probability of C, given the assumption of M, as conditioned by his or her knowledge, K.[6] In these terms, it is straightforward to restate $AR = OR \times DR$ as:

$$P_K(M \,\&\, -C \,\&\, -D) = P_K(M \,\&\, -C) \times P_K(-D \mid M \,\&\, -C). \tag{1}$$

It is not straightforward, however, to restate $OR = IR \times CR$. Consider equation (2):

$$P_K(M \,\&\, -C) = P_K(M) \times P_K(-C \mid M). \tag{2}$$

Although theoretically sound, equation (2) is inconsistent with IR and CR as defined in the professional standards. The term $P_K(-C \mid M)$ reflects the detective effect of controls, but not their preventive effect. Also, $P_K(M)$ fails to reflect the assumption that there are no related controls. Moreover, since the auditor normally knows something about the auditee's controls, $P_K(M)$ may entail rejection of the assumption underlying IR. Similar inconsistencies arise for a decomposition based on the event structure in figure 2 (see the appendix).

The decomposition in equation (2) is theoretically sound, in part because its event structure involves a clear sequence. In contrast, the audit risk model fails to adopt a clear sequence for prevention and detection by controls vis-à-vis the occurrence of a misstatement. The definition of CR, by encompassing both the preventive and detective aspects of controls, confuses their temporal relation with misstatements. By assuming no related controls, the definition of IR fails to recognize the possibility that the preventive effect of controls precedes the (non)occurrence of misstatements. Because of its ill-defined event structure, the audit risk model's decomposition of OR into IR and CR is inconsistent with theory-based decompositions. Indeed, when auditors assess IR and CR in accord with the professional standards, there is no logical basis for concluding that $OR = IR \times CR$ or $AR = DR/(IR \times CR)$, and auditors may inadvertently fail to plan audits in line with their target AR.

As a practical matter, it is possible that auditors do not adhere strictly to professional standards when assessing IR and CR; they may find it natural to employ an intuitive version of equation (2). Under this approach, IR and CR assessments are responses to the following queries: On the basis of prior knowledge and available information (including that about auditee controls), what is the risk of misstatement prior to the application of detective controls, and, if a misstatement has occurred, what is the risk that detective controls are ineffective? This approach may introduce a knowledge-based dependence between an auditor's *IR* and *CR* assessments, a dependence that stems from prior knowledge about causal relations among auditee conditions, controls, and misstatements, and from information about risk factors observed in the audit circumstances. Equation (2) accommodates this possibility by conditioning the auditor's assessments on *K*. In contrast, the audit risk model does not recognize such conditionality or dependence although professional standards emphasize the need to consider various risk factors.

---

[5] In addition to the principles of decomposition discussed earlier, decision analysts stress the following points regarding probability assessment (see Von Winterfeldt and Edwards 1986, chap. 4). First, the event structure must be clearly specified for the assessor. Second, uncertainty is a property of the assessor's knowledge about an event, rather than of the event itself, which supports the subjective interpretation of probability. Third, all subjective probabilities are conditional, whether or not the conditioning events or information are explicit. The probabilistic view of audit risk in the text incorporates each point.

[6] Though cumbersome, this notation is useful in distinguishing conditionality with respect to current knowledge and conditionality with respect to an assumed condition or information yet to be added (Hacking 1967).

Whether a knowledge-based dependence would result in a positive or negative association between auditors' *IR* and *CR* assessments in field settings is not clear *a priori*. Cushing and Loebbecke (1983) propose that when an auditee's controls are weak, its employees are more prone to commit intentional and unintentional errors. Thus, an auditor who perceives weak controls would assess both *CR* and *IR* as high because weak controls induce error-prone conditions, other things being equal, indicating a positive association between *IR* and *CR*. From another viewpoint, Graham (1985) observes that the auditee management typically sets up stronger controls over risky financial statement items. Thus, an auditor who identifies a risky financial statement item (e.g., inventory of precious metals) would assess *IR* as high, but would expect the quality of related controls to be high and *CR* to be low, other things equal, indicating a negative association between *IR* and *CR*. Another possible source of dependence is overlap in the observed specific or general risk factors used to assess *IR* and *CR*. Specific factors include an asset's susceptibility to theft (relevant to *IR*) and the separation of duties for the asset's handling and recordkeeping (relevant to *CR*). General factors include management's philosophy, operating style, and personnel policies (relevant to both *IR* and *CR*). If auditors weight general factors more heavily than specific factors, then a positive association between *IR* and *CR* is likely. In sum, *a priori* analyses support the prediction of *some* knowledge-based dependence between *IR* and *CR*, but yield conflicting predictions as to whether the association is positive or negative, suggesting Proposition 1:

Proposition 1: There is an association between auditors' *IR* and *CR* assessments.

The association between *IR* and *CR* may be affected by factors other than a knowledge-based dependence. Professional standards state that auditors should, under certain conditions, incorporate cost considerations into their risk assessments. When an auditor believes that the cost of making risk assessments, in terms of the time and effort spent on information processing, outweighs the likely benefit, in terms of a higher *DR*, audit procedures should be selected as if risk were maximal. Thus, high risk assessments may represent either auditors' beliefs about risk or cost considerations.

Misstatements and Risk Assessments at the Assertion Level

Another dimension of the audit risk model's decomposition involves the focus of risk assessment at the assertion versus account level. Over the last decade, the professional standards have shifted the focus by requiring application of the audit risk model, including *IR* and *CR* assessments, at the assertion level. This policy has the potential advantage of enhancing audit effectiveness and efficiency through more finely tuned risk assessments and test procedures. However, questions arise with respect to dependencies in risk assessments and test procedures over assertions. For example, Kinney (1992) examines analytically the adverse consequences of the model's failure to reflect the risk of misstatements that occur jointly in two or more assertions.

Viewed generally over many audit engagements, a policy of assertion-level risk assessments depends on several premises, including: (1) varying the rate of misstatements over assertions for an account; otherwise, any gain in effectiveness from extending the decomposition to the assertion level would be much lower; (2) varying auditors' risk assessments over assertions for an account; otherwise, their assessments would be informationally redundant; (3) positive association between the rate of misstatements and auditors' risk assessments; otherwise, auditors' ability to perform the task would be in doubt. Empirical propositions relating to these premises are stated below.

Rate of Misstatements Over Assertions. In conformance with equation (1), OR may be represented as the auditor's subjective probability, $P_K(M')$, where M' designates the conjunction of M and $-C$. Objectively, OR is more difficult to define in a given audit engagement (see Shafer and Srivastava [1990] for discussion). Over many engagements, however, OR is objectively reflected by the rate of misstatements for an assertion. Let $R(M')$ designate the rate of misstatements, which are not detected by controls, for an assertion over a large number of engagements. Then, the first premise of the assertion-level policy is that $R(M')$ varies over assertions for an account.

Testing whether $R(M')$ varies over assertions is complicated by two related problems. First, with few exceptions, misstatements are observable only when detected by the auditor. In conformance with equation (1), let D designate the auditor's detection of a misstatement in the assertion of interest. The rate of *detected* misstatements, $R(M' \,\&\, D)$, is computable from archival data but $R(M')$ is not. This problem precludes a direct test of the first premise. Second, the use of $R(M' \,\&\, D)$ as a surrogate for $R(M')$ is confounded by the link between an assessment of OR and the choice of test procedures. That is, the relation between $R(M')$ and $R(M' \,\&\, D)$ may be stated as:

$$R(M') = \frac{R(M' \,\&\, D)}{R(D|M')}. \tag{3}$$

Suppose $R(M' \,\&\, D)$ is observed to be higher for one assertion than another for a given account. The conclusion that $R(M')$ is proportionately higher would be valid if $R(D|M')$ is the same for each assertion, a premise that may not be true. A prior belief that OR is higher would lead to a lower DR, which implies the application of more powerful test procedures and a higher $R(D|M')$. In light of these problems, Proposition 2 is offered.

> *Proposition 2*: The rate of detected misstatements varies over assertions, after controlling for DR.

Auditors' Risk Assessments Over Assertions. Although assessing risk for each assertion would imply variability in such assessments, there are reasons to expect auditors to make similar risk assessments over assertions for an account. A knowledge-based dependence may exist among risk assessments for multiple assertions. Professional standards state that risk factors may be assertion-specific or general in their effects. If auditors weight general factors more heavily than assertion-specific factors, an association in risk assessments over assertions would be likely. Also, the auditor's varying professional experience may produce association between different risk assessments. Furthermore, auditors often use test procedures that provide evidence relevant to multiple assertions. Anticipating usage of such procedures, auditors may allow the risk assessment for one assertion to determine the risk assessment for another. In a like manner, when assessing risk for multiple assertions for an account, an auditor may simplify the task by employing a heuristic that identifies the "most important" assertion (e.g., one linked to overstatement in the case of assets or understatement in the case of liabilities), deliberately assess *IR* and *CR* for this assertion, and mechanically assess the same risk for other assertions. Finally, some auditors may have difficulty in discriminating risk beyond the account level, especially under an audit's time constraint. For these reasons, the association in auditors' risk assessments over assertions is expected to be nontrivial, but less than perfect. Proposition 3 reflects the policy requirement.

Proposition 3: Auditors' *IR* and *CR* assessments vary over assertions for an account.

Risk Assessment Accuracy. Auditors' effectiveness in risk assessment depends, in part, on accuracy with respect to the events that occur. Typically, evaluation of individuals' risk assessment accuracy in behavioral decision research (Lichtenstein et al. 1982; Von Winterfeldt and Edwards 1986; see Solomon and Shields [1993] for a review of experimental auditing studies) is done at an aggregate level (over many individuals and events) and relies on the notion of *calibration*. To explain, suppose a firm's auditors assess *OR* in the form of subjective probability, $P_K(M')$, on each of many audit engagements. Calibration requires, for each subset of assessments in which $P_K(M')$ has the same value, $P_K(M')$ must be equal to $R(M')$, the rate of occurrence for M' in that subset. When *OR* assessments are made on an ordinal scale, calibration requires the association between *OR* and $R(M')$ to be positive. However, because $R(M' \ \& \ D)$ is known, but $R(M')$ is not, *DR* is included as a control variable. Also, since there is no logical basis for concluding that $OR = IR \times CR$, the focus is on the association between *IR* and $R(M' \ \& \ D)$, after controlling for *CR* (as well as *DR*).[7] Incorporating these controls, Proposition 4 addresses the accuracy issue.

Proposition 4: There is a positive association between auditor's; *IR* assessments and the rate of detected misstatements, after controlling for *CR* and *DR*.

II. Data Analysis

The data were collected as part of a project conducted by KPMG Peat Marwick's internal research group. A random sample of 385 engagements was selected from the firm's audit billing file. A questionnaire, under the signature of the firm partner heading the research group, was sent to the auditor in charge of each engagement. Among other things, information from the audit workpapers was requested on three variables pertaining to each of five assertions (see fn. 1) for trade accounts receivable, inventory, and trade accounts payable. One variable is whether the auditor detected a misstatement(s) exceeding tolerable error (yes, no). The other variables are about the *IR* assessment (low/moderate/high) and *CR* assessment (very low/low/moderate/high). For

---

[7] Although there is no logical basis for concluding that $OR = IR \times CR$, it is reasonable to suppose that *OR* increases with *IR*, when holding *CR* constant. The test of Proposition 4 relies on the latter relation.

Table 1

Auditees by Industry and Account

Industry	*Trade Accounts Receivable*	*Inventory*	*Trade Accounts Payable*
Agribusiness	9	8	9
Banking	0	0	7
Educational and other institutions	20	6	26
Government	8	5	11
Health care	14	9	14
High technology	11	10	11
Insurance	8	0	5
Manufacturing	39	40	41
Merchandising	13	13	12
Real estate	7	2	10
Thrift	0	0	3
Other	52	30	60
Total	181	123	209

the 385 engagements, 254 questionnaires were returned. There were 82 less than full-scope audit engagements and 49 nonresponses. The respondents also were asked, for comparison purposes, to complete a questionnaire for an engagement not involving an EDP system; 20 such questionnaires were returned, increasing the total sample to 274. Fifty-nine of the 274 engagements did not require risk assessments for any of the three accounts. Of the remaining 215 engagements, 209 required risk assessments for trade accounts payable, 181 for trade accounts receivable, and 123 for inventory. The data from these engagements are examined in this section.[8]

Table 1 shows the auditees' industry classification, showing manufacturing to comprise a sizeable proportion of the sample. Separate analyses are reported overall and for the manufacturing subsample. The balance sheet date falls in 1988 for 171 auditees, 1989 for 37, 1987 for six, and 1986 for one. The median value (interquartile range) is $15,480,000 ($3,941,000–$55,192,000) for total assets, $14,444,000 ($2,731,000–$50,124,000) for total revenue, and $702,000 ($142,000–$2,768,000) for net income.

Association Between IR *and* CR *Assessments*

Tables 2 and 3 present the joint distributions of *IR* and *CR* by account and assertion, overall and for the manufacturing subsample, respectively. Collapsing over accounts and assertions, the modal value is moderate for *IR* (58.8 percent overall, 61.8 percent for manufacturing auditees) and high for *CR* (81.3 percent overall, 82.5 percent for

---

[8] There are missing data in a few cases, which explains the slight variation in the number of observations in various analyses.

Table 2
Inherent Risk (*IR*) and Control Risk (*CR*) by Assertion and Account—All Auditees

	Trade Accounts Receivable			*Inventory*			*Trade Accounts Payable*		
	IR_h	IR_m	IR_l	IR_h	IR_m	IR_l	IR_h	IR_m	IR_l
Completeness:									
CR_h	10	87	44	17	42	40	8	111	48
CR_m	0	20	2	2	11	2	1	19	6
CR_l	1	5	2	0	4	2	0	5	5
CR_{vl}	0	1	1	0	0	1	0	1	1
$T(p)$*		-0.05 (0.12)			-0.04 (0.27)			0.03 (0.26)	
$T(p)$**		0.09 (0.18)			0.27 (0.03)			0.19 (0.05)	
Existence:									
CR_h	9	91	45	16	38	42	6	93	65
CR_m	1	18	2	2	11	2	1	19	6
CR_l	1	6	2	0	6	3	0	5	7
CR_{vl}	0	1	1	0	0	1	0	1	1
$T(p)$*		-0.06 (0.07)			-0.05 (0.22)			-0.01 (0.42)	
$T(p)$**		0.11 (0.16)			0.28 (0.03)			0.24 (0.02)	
Accuracy:									
CR_h	13	92	39	18	46	34	7	104	55
CR_m	2	19	1	2	11	2	1	19	6
CR_l	1	5	2	0	5	2	0	6	6
CR_{vl}	0	1	1	0	0	1	0	1	1
$T(p)$*		-0.06 (0.10)			-0.01 (0.43)			0.01 (0.40)	
$T(p)$**		0.16 (0.07)			0.25 (0.04)			0.20 (0.04)	
Valuation:									
CR_h	23	98	26	28	41	33	6	91	66
CR_m	2	16	1	1	9	1	1	17	7
CR_l	1	5	3	1	4	2	0	6	6
CR_{vl}	0	1	1	0	0	1	0	2	1
$T(p)$*		0.02 (0.29)			0.02 (0.37)			-0.02 (0.35)	
$T(p)$**		0.22 (0.04)			0.21 (0.11)			0.14 (0.12)	
Ownership:									
CR_h	8	86	50	16	40	41	6	91	68
CR_m	0	18	2	2	11	2	1	17	6
CR_l	1	6	3	0	5	3	0	5	9
CR_{vl}	0	1	1	0	0	1	0	1	1
$T(p)$*		-0.06 (0.10)			-0.03 (0.28)			0.01 (0.42)	
$T(p)$**		0.13 (0.13)			0.30 (0.02)			0.27 (0.01)	

* Kendall's T (significance level), including all data in joint distribution.

** Kendall's T (significance level), excluding CR_h data.

manufacturing auditees). The combination of IR_m/CR_h was assessed 45.9 percent of the time overall and 50.4 percent of the time for manufacturing auditees.

The association between *IR* and *CR* was examined for each account-assertion pair

Table 3
Inherent Risk (*IR*) and Control Risk (*CR*) by Assertion and Account—Manufacturing Auditees

	Trade Accounts Receivable			*Inventory*			*Trade Accounts Payable*		
	IR_h	IR_m	IR_l	IR_h	IR_m	IR_l	IR_h	IR_m	IR_l
Completeness:									
CR_h	1	21	9	6	20	7	2	23	10
CR_m	0	5	1	1	2	1	0	3	0
CR_l	0	1	0	0	1	0	0	1	1
CR_{vl}	0	0	1	0	0	1	0	0	1
$T(p)$	0.01 (0.49)			0.05 (0.28)			0.04 (0.29)		
Existence:									
CR_h	1	21	9	5	19	8	2	17	15
CR_m	0	5	1	1	2	1	0	3	0
CR_l	0	1	0	0	2	0	0	1	2
CR_{vl}	0	0	1	0	0	1	0	0	1
$T(p)$	0.01 (0.49)			0.03 (0.39)			0.03 (0.38)		
Accuracy:									
CR_h	1	22	8	5	23	4	2	20	12
CR_m	0	6	0	1	2	1	0	3	0
CR_l	0	1	0	0	2	0	0	1	2
CR_{vl}	0	0	1	0	0	1	0	0	1
$T(p)$	-0.04 (0.33)			0.07 (0.20)			0.06 (0.24)		
Valuation:									
CR_h	5	22	5	10	21	4	2	15	17
CR_m	0	5	0	1	1	0	0	2	0
CR_l	0	1	0	0	1	0	0	1	2
CR_{vl}	0	0	1	0	0	1	0	1	1
$T(p)$	0.07 (0.22)			0.04 (0.28)			-0.01 (0.47)		
Ownership:									
CR_h	1	20	10	4	20	8	2	16	16
CR_m	0	5	1	1	2	1	0	3	0
CR_l	0	1	0	0	2	0	0	1	2
CR_{vl}	0	0	1	0	0	1	0	0	1
$T(p)$	-0.01 (0.44)			0.02 (0.42)			0.01 (0.44)		

by using the Kendall rank-order correlation coefficient, T.[9] The results for each pair are shown below the corresponding joint distribution in tables 2 and 3. For example, for completeness of trade accounts payable in table 2, $t=0.03$ ($p=0.26$), which indicates an insignificant association between *IR* and *CR* for this account-assertion pair. Overall and for the manufacturing subsample, *t* does not differ significantly from zero for any account-assertion pair. These results indicate a negligible association between *IR* and *CR* and fail to support Proposition 1.

---

[9] Additional analyses with other nonparametric measures of association yielded very similar results. All nonparametric tests used in this study are described in Siegel and Castellan (1988).

It is important to consider the highly skewed *CR* distribution: more than 80 percent are CR_h. As stated earlier, a CR_h assessment may not reflect a belief that controls are ineffective but rather a decision not to rely on controls for efficiency reasons, which seems to occur more frequently. Even though the overall association between *IR* and *CR* is negligible, a knowledge-based dependence may affect the association in the absence of a decision not to rely on controls. This possibility can be pursued only in a limited way with the archival data by examining those cases in which some reliance was placed on controls (i.e., when *CR* was assessed below the maximum). The results denoted by two asterisks in table 2 are for these cases.[10] In contrast with the overall analysis, *t*-statistics are consistently positive and are significantly different from zero for nine of the 15 pairs of account-assertion. This provides some support for a positive, albeit low to moderate, association between *IR* and *CR*.

Rate of Detected Misstatements

Columns I, III, and V of table 4 show the rate of detected misstatements, $R(M' \& D)$, by account and assertion. For example, out of 209 cases involving completeness of trade accounts payable, a misstatement exceeding tolerable error was detected in 24 cases, such that $R(M' \& D)=0.115$. The difference between each assertion's $R(M' \& D)$ and the mean over assertions is shown in parentheses. Including all auditees, $R(M' \& D)$ varies significantly over assertions for each account, when using the Cochran Q-test for related samples. For trade accounts receivable and inventory, $R(M' \& D)$ is highest for valuation and lowest for ownership. For trade accounts payable, $R(M' \& D)$ is highest for completeness and lowest for ownership. Including manufacturing auditees only, $R(M' \& D)$ varies significantly over assertions for inventory, but not for trade accounts receivable or trade accounts payable. The insignificant results for the latter accounts are attributable to few misstatements being detected for any assertion, as well as to the limited number of observations for manufacturing auditees. Prior to considering the effect of *DR*, the overall results support Proposition 2.

Regarding the effect of *DR*, an additional analysis was performed in which $R(M' \& D)$ was divided by 1.00 minus a weighted-average *DR*. For each account-assertion pair, a weighted-average *DR* was computed by using KPMG Peat Marwick's audit manual which relates each possible combination of *IR*, *CR*, and *AP* to a numerical *TD* value, assuming a target *AR* of 0.04 (see fn. 3 for definitions of *AP* and *TD*). Table 5 presents an excerpt of the *TD* values when *AP* is high.[11] Because $AR=0.04$, the first entry of $TD=0.04$ implies that $AP_h=1.00$, and $DR=TD\times AP_h=TD$ for all entries. Thus, table 5 shows *DR* as it corresponds to each possible combination of *IR* and *CR*, and the table may be used to compute a weighted-average *DR* for any subset of the archival data. To illustrate, consider the joint distribution of *IR* and *CR* for completeness of trade accounts payable (table 2). According to table 5, *DR* is 0.04 in the eight cases with IR_h/CR_h, 0.06 in the 111 cases with IR_m/CR_h, and so on, which yields a weighted-average *DR* of 0.11 for this account-assertion pair. Subtracting the weighted-average *DR*

---

[10] Data limitations prevent a similar analysis for the manufacturing subsample.

[11] Though obviously not reflective of any given case, the values in table 5 presumably indicate the pattern expected by firm policymakers over many applications of the audit risk model.

Table 4
Rate of Detected Misstatements

	Trade Accounts Receivable		*Inventory*		*Trade Accounts Payable*	
	I	*II*	*III*	*IV*	*V*	*VI*
All Auditees:						
Completeness	0.044	0.049	0.081	0.091	0.115	0.130
	(-0.018)	(-0.020)	(0.000)	(0.000)	(+0.059)	(+0.066)
Existence	0.044	0.049	0.049	0.056	0.034	0.039
	(-0.018)	(-0.020)	(-0.033)	(-0.036)	(-0.022)	(-0.025)
Accuracy	0.089	0.099	0.114	0.128	0.072	0.082
	(+0.027)	(+0.030)	(+0.033)	(+0.037)	(+0.016)	(+0.018)
Valuation	0.127	0.141	0.154	0.172	0.034	0.039
	(+0.065)	(+0.072)	(+0.073)	(+0.081)	(-0.022)	(-0.024)
Ownership	0.006	0.007	0.008	0.009	0.024	0.027
	(-0.057)	(-0.063)	(-0.073)	(-0.082)	(-0.032)	(-0.036)
Q	37.74		30.31		39.25	
(*p*)	(0.001)		(0.001)		(0.001)	
Manufacturing Auditees:						
Completeness	0.026	0.030	0.100	0.112	0.049	0.056
	(+0.010)	(+0.012)	(+0.015)	(+0.017)	(+0.010)	(+0.010)
Existence	0.000	0.000	0.025	0.028	0.049	0.057
	(-0.015)	(-0.017)	(-0.060)	(-0.067)	(+0.010)	(+0.012)
Accuracy	0.000	0.000	0.125	0.141	0.073	0.085
	(-0.015)	(-0.017)	(+0.040)	(+0.046)	(+0.034)	(+0.040)
Valuation	0.051	0.057	0.175	0.193	0.024	0.028
	(+0.036)	(+0.040)	(+0.090)	(+0.098)	(-0.015)	(-0.016)
Ownership	0.000	0.000	0.000	0.000	0.000	0.000
	(-0.015)	(-0.017)	(-0.085)	(-0.095)	(-0.039)	(-0.045)
Q	13.83		5.33		4.33	
(*p*)	(0.01)		(0.25)		(0.36)	

Note: Columns I, III, and V show $R(M' \& D)$ by assertion and, in parentheses, its difference from the mean over assertions for the account. Columns II, IV, and VI show $R(M' \& D)/[1\text{-}DR]$ by assertion and, in parentheses, its difference from the mean over assertions for the account.

from 1.00 provides an indicator of $R(D|M')$ in equation (3), and dividing $R(M' \& D)$ by the indicator provides an estimate of $R(M')$. This procedure was applied to each account-assertion pair, overall and for manufacturing auditees. Columns II, IV, and VI of table 4 contain the results. Although the $R(M')$ estimates are, of course, higher than $R(M' \& D)$, the difference between each assertion's $R(M')$ estimate and the mean over assertions is very close to the corresponding difference shown in column I, III, or V.[12] This is because the similarity in the joint distributions of *IR* and *CR* over assertions (see

---

[12] The Cochran *Q*-test cannot be repeated on the rates of detected misstatements adjusted for *DR*.

Table 5
Detection Risk (*DR*) Values

Combinations of Inherent and Control Risk	*DR = TD*
IR_h/CR_h	0.04
IR_m/CR_h	0.06
IR_l/CR_h	0.11
IR_h/CR_m	0.09
IR_m/CR_m	0.15
IR_l/CR_m	0.25
IR_h/CR_l	0.22
IR_m/CR_l	0.37
IR_l/CR_l	0.61
IR_h/CR_{vl}	0.37
IR_m/CR_{vl}	0.61
IR_l/CR_{vl}	1.00

Note: Entries are test of details (*TD*) risk, given analytical procedures risk (*AP*) is high (taken from the KPMG Peat Marwick audit manual with permission).

Table 6
Co-occurrences of Detected Misstatements in Assertions

Misstated Assertion(s)	*Trade Accounts Receivable*	*Inventory*	*Trade Accounts Payable*
Completeness (C)	2	3	11
Existence (E)	2	1	1
Accuracy (A)	7	1	4
Valuation (V)	13	6	0
Ownership (O)	0	0	1
	24	11	17
CE	0	0	1
CA	0	1	5
CV	1	0	2
EA	1	0	0
EV	1	1	0
AV	2	6	1
CEA	0	1	1
CEV	0	1	0
CAV	2	3	0
EAV	1	1	0
CEAV	2	0	0
CEAVO	1	1	4
	11	15	14
Total	35	26	31

below for further analysis) produces a fairly constant weighted-average *DR* over assertions. In short, no difference in *OR* implies no difference in *DR*. The test of Proposition 2 is minimally affected by variation in *DR*.

Table 6 provides another view of the relation in risk over assertions by showing the co-occurrences of detected misstatements in different assertions for the same

Table 7
Risk Assessments Over Assertions

	Number of Assertions With Same Assessment	*Trade Accounts Receivable*	*Inventory*	*Trade Accounts Payable*
All Auditees				
IR	All five	133 (76.5%)	94 (77.7%)	178 (87.3%)
	Four	23 (13.2%)	14 (11.6%)	9 (4.4%)
	Three	18 (10.3%)	12 (9.9%)	16 (7.8%)
	Two	—	1 (0.8%)	1 (0.5%)
		174 (100.0%)	121 (100.0%)	204 (100.0%)
	X^2 (p)*	9.08 (0.06)	6.79 (0.14)	5.47 (0.24)
CR	All five	173 (97.2%)	115 (93.5%)	199 (97.5%)
	Four	3 (1.7%)	5 (4.1%)	2 (1.0%)
	Three	2 (1.1%)	3 (2.4%)	3 (1.5%)
		178 (100.0%)	123 (100.0%)	204 (100.0%)
	X^2 (p)*	0.07 (0.99)	0.46 (0.97)	0.15 (0.99)
Manufacturing Auditees				
IR	All five	29 (74.3%)	23 (60.5%)	33 (80.4%)
	Four	6 (15.4%)	8 (21.1%)	4 (9.8%)
	Three	4 (10.3%)	7 (18.4%)	4 (9.8%)
		39 (100.0%)	38 (100.0%)	41 (100.0%)
	X^2 (p)*	2.77 (0.59)	5.03 (0.28)	2.61 (0.62)
CR	All five	38 (97.4%)	38 (97.4%)	39 (95.1%)
	Four	1 (2.6%)	1 (2.6%)	2 (4.9%)
		39 (100.0%)	39 (100.0%)	41 (100.0%)
	X^2 (p)*	0.05 (0.99)	0.43 (0.98)	0.12 (0.99)

* Results of Friedman test for differences in risk assessments over assertions for an account.

account.[13] With the assumption that any assertion-level misstatement implies an account-level misstatement, the rate of detected misstatements is 0.194 for trade accounts receivable, 0.211 for inventory, and 0.150 for trade accounts payable. Each account-level rate is considerably lower than the sum of corresponding assertion-level rates (trade accounts receivable, 0.310; inventory, 0.406; trade accounts payable, 0.279), because of the joint occurrence of misstatements in multiple assertions. Two or more assertions are implicated in 43.5 percent of the cases with an account-level misstatement. These results suggest that decomposing risk to the assertion level should accommodate the dependence in risk over assertions (cf. Kinney 1992).

Association in Risk Assessments Over Assertions

Table 7 presents evidence on the pattern of *IR* and *CR* assessments over assertions for each account. Including all auditees, differences in *IR* over assertions are marginally significant for trade accounts receivable, but insignificant for inventory and trade accounts payable, when using the Friedman test for related samples. For trade accounts receivable, *IR* is somewhat higher for valuation versus other assertions (table 2). Differences in *CR* over assertions are insignificant for each account. Table 7 also shows the frequencies with which auditors made the same assessment for all five (or fewer) assertions. For example, focusing on trade accounts payable, auditors assessed the same *IR* for all five assertions 87.3 percent of the time and the same *CR* 97.5 percent of the time. Including manufacturing auditees only, differences in *IR* and *CR* over assertions are insignificant for each account, and the tendency is to assess the same *IR* and *CR* over assertions. In sum, contrary to Proposition 3, the evidence indicates that auditors generally do not vary their risk assessments over assertions. The *CR* pattern may be explained by a predominance of decisions not to rely on controls, which imply CR_h for all assertions. However, in light of the high frequencies of IR_l and IR_m, the *IR* pattern must be explained along different lines, such as use of the "most important" assertion heuristic described earlier.

---

[13] For a given audit engagement, misstatements in two or more assertions do not necessarily result from the same transaction or event.

IR *Assessment Accuracy*

Table 8 shows $R(M' \,\&\, D)$ for each subset of *IR* assessments with the same value (low/moderate/high) for each account-assertion pair. The table also shows the corresponding $R(M')$ estimates in parentheses, which were computed by dividing 1.00 minus *DR* into $R(M' \,\&\, D)$, where *DR* equals 0.04 for IR_h, 0.06 for IR_m, and 0.11 for IR_l (see table 5). For example, focusing on the subset of engagements with an assessment of IR_l for completeness of trade accounts payable, $R(M' \,\&\, D)$ is 0.083, and the $R(M')$ estimate is 0.094. The last column reports Kendall's *T*, which indicates the degree of association between *IR* and $R(M' \,\&\, D)$. To control for *CR*, the table includes only cases involving CR_h.[14]

Table 8 also reveals that $R(M' \,\&\, D)$ increases monotonically with *IR* for only seven of the 15 account-assertion pairs. For all eight pairs with inconsistencies, $R(M' \,\&\, D)$ is not higher for IR_h versus IR_m, and for one pair, $R(M' \,\&\, D)$ is not higher for IR_m versus IR_l. Because of the low number of observations for IR_h, it is reasonable to consider monotonicity after combining the IR_h and IR_m subsets. With this re-grouping, monotonicity holds for 13 of 15 pairs. Using a binomial test, the probability of randomly observing 13 of 15 cases with the same sign is less than 0.01. Both exceptions are minimal in the sense that monotonicity would result by re-classifying one case with a detected misstatement from the IR_l subset to the $IR_{h,m}$ subset. In addition, *T* generally indicates a low, positive association between *IR* and $R(M' \,\&\, D)$, which is statistically significant only for accuracy and valuation of inventory (marginally significant for completeness of inventory and trade accounts payable). These results suggest that auditors' *IR* assessments are weakly consistent with calibration, which provides a modest support for Proposition 4.

---

[14] Given CR_h, *DR* ranges only from 0.04 for IR_h to 0.11 for IR_l (table 5). This range leads to small differences between $R(M' \,\&\, D)$ and the $R(M')$ estimate. Consequently, the evaluation of *IR* assessment accuracy is minimally affected by using $R(M' \,\&\, D)$ as a surrogate for $R(M')$. Data limitations prevent a similar evaluation for the manufacturing subsample.

Table 8
Inherent Risk (IR) Assessment Accuracy

	IR_h	IR_m	IR_l	T (p)
Trade Accounts Receivable				
Completeness	1/10 0.100(0.104)	5/87 0.057(0.061)	1/44 0.023(0.026)	0.041 (0.13)
Existence	0/9 0.000(0.000)	7/91 0.077(0.082)	0/45 0.000(0.000)	0.048 (0.09)
Accuracy	1/13 0.077(0.080)	12/92 0.130(0.139)	2/39 0.051(0.058)	0.045 (0.18)
Valuation	4/23 0.174(0.181)	13/98 0.133(0.141)	3/26 0.115(0.130)	0.032 (0.28)
Ownership	0/8 0.000(0.000)	1/86 0.012(0.012)	0/50 0.000(0.000)	0.008 (0.28)
Inventory				
Completeness	2/17 0.118(0.123)	5/42 0.119(0.127)	1/40 0.025(0.028)	0.090 (0.06)
Existence	1/16 0.063(0.065)	3/38 0.079(0.084)	1/42 0.024(0.027)	0.045 (0.18)
Accuracy	4/18 0.222(0.231)	5/46 0.109(0.116)	1/34 0.029(0.033)	0.140 (0.02)
Valuation	8/28 0.286(0.298)	7/41 0.171(0.182)	1/33 0.030(0.034)	0.215 (0.01)
Ownership	0/16 0.000(0.000)	0/40 0.000(0.000)	1/41 0.024(0.027)	-0.024 (0.14)
Trade Accounts Payable				
Completeness	3/8 0.375(0.391)	13/111 0.117(0.125)	4/48 0.083(0.094)	0.075 (0.06)
Existence	0/6 0.000(0.000)	3/93 0.032(0.034)	2/65 0.031(0.035)	-0.003 (0.45)
Accuracy	1/7 0.143(0.149)	8/104 0.077(0.082)	2/55 0.036(0.041)	0.047 (0.11)
Valuation	1/6 0.167(0.174)	3/91 0.033(0.035)	2/65 0.031(0.035)	0.021 (0.24)
Ownership	0/6 0.000(0.000)	2/91 0.022(0.023)	1/67 0.015(0.017)	0.004 (0.43)

Note: Entries in the first three columns are $R(M' \,\&\, D)$ and, in parentheses, $R(M' \,\&\, D)/[1\text{-}DR]$ where DR equals 0.04 for IR_h, 0.06 for IR_m, and 0.11 for IR_l. Entries in the last column are Kendall's T and its significance level. The table includes cases with CR_h only.

These results are consistent with auditors' use of the "most important" assertion heuristic, suggesting that the analysis in table 8 may understate the extent of empirical support for Proposition 4 by including all account-assertion pairs instead of focusing

Table 9
Inherent Risk (*IR*) Assessment Accuracy for "Most Important" Assertions

Possible Combinations of "Most Important" Assertions						
Trade Accounts Receivable	*Inventory*	*Trade Accounts Payable*	IR_h	IR_m	IR_l	T (p)
Valuation	Valuation	Completeness	15/59 0.254(0.265)	33/250 0.132(0.140)	8/107 0.075(0.084)	0.103 (0.001)
Accuracy	Valuation	Completeness	12/49 0.245(0.255)	32/244 0.131(0.140)	7/120 0.058(0.066)	0.108 (0.001)
Valuation	Accuracy	Completeness	11/49 0.224(0.234)	31/255 0.122(0.129)	8/108 0.074(0.083)	0.080 (0.006)
Accuracy	Accuracy	Completeness	8/39 0.205(0.214)	30/249 0.120(0.128)	7/121 0.058(0.065)	0.081 (0.004)
Valuation	Valuation	Accuracy	13/58 0.224(0.233)	28/243 0.115(0.123)	6/114 0.053(0.059)	0.102 (0.001)
Accuracy	Valuation	Accuracy	10/48 0.208(0.217)	27/237 0.114(0.121)	5/127 0.039(0.044)	0.102 (0.001)
Valuation	Accuracy	Accuracy	9/48 0.188(0.195)	26/248 0.105(0.112)	6/115 0.052(0.059)	0.077 (0.005)
Accuracy	Accuracy	Accuracy	6/38 0.158(0.164)	25/242 0.103(0.110)	5/128 0.039(0.044)	0.074 (0.005)

Note: Entries in the first three columns are $R(M' \& D)$ and, in parentheses, $R(M' \& D)/[1\text{-}DR]$ where DR equals 0.04 for IR_h, 0.06 for IR_m, and 0.11 for IR_l. Entries in the last column are Kendall's T and its significance level. The table includes cases with CR_h only.

on the "most important" assertion for each account. Further, because of the generally low $R(M' \& D)$, regardless of IR, along with the limited number of observations per account-assertion pair, each test at the disaggregated level lacks power. In light of these considerations, a final analysis was done at an aggregated level. The analysis examines various combinations of "most important" assertions for the three accounts (table 9).[15] For each combination, the data pertaining to three account-assertion pairs were aggregated, and the analysis in table 8 was repeated. The results for each combination are similar. Relative to the results in table 8, t-statistics generally are somewhat higher and significant. More striking perhaps is the stable pattern of $R(M' \& D)$ over IR levels, which is consistent with calibration. Assuming that auditors use the heuristic in line with any of the combinations in table 9, the evidence supports Proposition 4 with respect to "most important" assertions.

---

[15] KPMG Peat Marwick's audit manual and audit program guide contain no reference to the "most important" assertion for a given account, and there is no prior research that is directly relevant. Thus, the combinations shown in table 9 are somewhat arbitrary. These assertions were selected because they reflect auditors' concern about overstatement of assets and understatement of liabilities, and because $R(M' \& D)$ is higher for these assertions versus others (table 4). For a complete analysis of all 125 possible combinations, contact the author.

III. Conclusion

The finding of an insignificant association between *IR* and *CR* runs counter to the idea that a knowledge-based dependence exists between the audit risk model's components. This result largely stems from the predominance of cases in which *CR* is assessed at the maximum. An explanation is that auditors use *CR* assessments primarily as a way to document decisions about (non)reliance on controls. Though consistent with professional standards, this practice creates doubt about whether the purpose of risk assessment is to guide or confirm the choice of audit procedures. A cost-based decision not to rely on controls implies CR_h and limits the effect of assessing *IR* to moderating *DR* between 0.04 and 0.11 (table 5). Under these conditions, the issue of whether a knowledge-based dependence underlies *IR* and *CR*, or whether the model's decomposition is theory-based, has limited decision significance. Relevant policy should have a logical foundation clarifying the event structure underlying the audit risk model in connection with its decomposition of *OR* into *IR* and *CR*.

This study's results suggest several questions for future research that might be best addressed through experimental methods. For example, auditors in experimental settings can be asked to provide risk assessments without regard to decisions about test procedures. In such settings, do *IR* and *CR* assessments covary? Is the association caused by a knowledge-based dependence? Experimentation also may be used to test alternative explanations of unvarying risk assessments over assertions. Is the pattern caused by using the "most important" assertion heuristic or by heavier weight on general versus assertion-specific risk factors? Finally, when experimental stimuli are based on actual audit engagements, it is possible to examine systematically the effects of task experience, knowledge, information, and formal aids on risk assessment accuracy. Such experiments would valuably complement this field-based research study.

Appendix

Figure 2 presents an event sequence that separates the preventive and detective aspects of controls. When formalized, the event structure has four components: C', the event of prevention by controls; M, the event of a misstatement; C'', the event of detection by controls; and D, the event of detection by the auditor. Thus, a misstatement in audited financial statements requires the conjunction of $-C'$, M, $-C''$, and $-D$. Given this event structure, it is straightforward to restate $AR = OR \times DR$ as:

$$P_K(-C' \,\&\, M \,\&\, -C'' \,\&\, -D) = P_K(-C' \,\&\, M \,\&\, -C'') \times P_K(-D \mid -C' \,\&\, M \,\&\, -C''). \tag{A1}$$

It is not straightforward, however, to restate $OR = IR \times CR$. Consider equation (A2):

$$\begin{aligned} P_K(-C' \,\&\, M \,\&\, -C'') &= P_K(-C' \,\&\, M) \times P_K(-C'' \mid -C' \,\&\, M), \\ &= P_K(M \mid -C') \times P_K(-C') \times P_K(-C'' \mid -C' \,\&\, M). \end{aligned} \tag{A2}$$

Equation (A2) is inconsistent with *IR* and *CR* as defined in the professional standards. $P_K(M \mid -C')$ assumes the condition of a prevention failure, which is not equivalent to the assumption of *IR* that there are no related controls. Also, unlike *CR*, equation (A2) contains two measures pertaining to the effectiveness of controls: $P_K(-C')$ concerns prevention and $P_K(-C'' \mid -C' \,\&\, M)$ concerns detection, and these do not generally combine into a single measure that corresponds to *CR*.

The professional standards define *CR* as the "risk that a material misstatement that could occur in an assertion *will not be prevented or detected* on a timely basis by an entity's internal control structure policies and procedures" (AICPA 1992, AU 319, italics added). The italicized phrase may be represented by the conjunction of $-C'$ and $-C''$, so that $P_K(-C' \& -C'')$ would be a single measure corresponding to *CR*:

$$\begin{aligned} P_K(-C' \& -C'') &= P_K(-C') \times P_K(-C'' \mid -C'), \\ &= P_K(-C') \times \{P_K(M \& -C'' \mid -C') + P_K(-M \& -C'' \mid -C')\}, \\ &= P_K(-C')\{P_K(-C'' \mid -C' \& M) \times P_K(M \mid -C') \\ &\quad + P_K(-C'' \mid -C' \& -M) \times P_K(-M \mid -C'')\}. \end{aligned} \tag{A3}$$

Since $P_K(-C'' \mid -C' \& -M) = 1$ and $P(-M \mid -C') = 1 - P(M \mid -C')$,

$$P_K(-C' \& -C'') = P_K(-C')\{[P_K(-C'' \mid -C' \& M) - 1]P_K(M \mid -C') + 1\}. \tag{A4}$$

To obtain consistency between equation (A4) and the last two terms of equation (A2), the former must reduce to:

$$P_K(-C' \& -C'') = P_K(-C') \times P_K(-C'' \mid -C' \& M). \tag{A5}$$

Equation (A5) holds only in the special case where $P_K(M \mid -C') = 1$ (i.e., a prevention failure necessitates the occurrence of a misstatement). However, a misstatement normally requires conditions beyond a prevention failure (e.g., employee action).

References

American Institute of Certified Public Accountants. 1992. *AICPA Professional Standards*, Vol. 1, *U.S. Auditing Standards*. New York: AICPA.

Colbert, J. L. 1988. Inherent risk: An investigation of auditors' judgments. *Accounting, Organizations and Society* 13 (January): 111–21.

Cushing, B. E., and J. K. Loebbecke. 1983. Analytical approaches to audit risk: A survey and analysis. *Auditing: A Journal of Practice & Theory* 3 (Fall): 23–41.

Daniel, S. J. 1988. Some empirical evidence about the assessment of audit risk in practice. *Auditing: Journal of Practice & Theory* 7 (Spring): 174–81.

Graham, L. E. 1985. Audit risk—Part I. *The CPA Journal* 55 (August): 12–21.

Hacking, I. 1967. Slightly more realistic personal probability. *Philosophy of Science* 34 (December): 311–25.

Jiambalvo, J., and W. Waller. 1984. Decomposition and assessments of audit risk. *Auditing: A Journal of Practice & Theory* 3 (Spring): 80–88.

Kinney, W. R. 1983. A note on compounding probabilities in auditing. *Auditing: A Journal of Practice & Theory* 3 (Spring): 13–22.

——. 1984. Discussant's response to "An analysis of the audit framework focusing on inherent risk and the role of statistical sampling in compliance testing." In *Auditing Symposium VII*, edited by H. F. Stettler and N. A. Ford. Lawrence, KS: Touche Ross/University of Kansas.

——. 1989. Achieved audit risk and the audit outcome space. *Auditing: A Journal of Practice & Theory* 8 (Supplement): 67–84.

——. 1992. The audit risk model at the financial statement level: The joint occurrence risk problem. Working paper, University of Texas.

Leslie, D. A. 1984. An analysis of the audit framework focusing on inherent risk and the role of statistical sampling in compliance testing. In *Auditing Symposium VII*, edited by H. F. Stettler and N. A. Ford. Lawrence, KS: Touche Ross/University of Kansas.

Libby, R., J. T. Artman, and J. J. Willingham. 1985. Process susceptibility, control risk, and audit planning. *The Accounting Review* 60 (April): 212–30.

Lichtenstein, S., B. Fischhoff, and L. D. Phillips. 1982. Calibration of probabilities: The state of the art to 1980. In *Judgment under Uncertainty: Heuristics and Biases*, edited by D. Kahneman et al. Cambridge, England: Cambridge University Press.

Raiffa, H. 1968. *Decision Analysis*. Reading, MA: Addison-Wesley.

Senetti, J. T. 1990. Toward a more consistent model for audit risk. *Auditing: A Journal of Practice & Theory* 9 (Spring): 103–12.

Shafer, G., and R. Srivastava. 1990. The Bayesian and belief-function formalisms—A general perspective for auditing. *Auditing: A Journal of Practice & Theory* 9 (Supplement): 110–37.

Siegel, S., and N. J. Castellan. 1988. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.

Solomon, I., and M. Shields. 1993. Judgment and decision research in auditing. In *Judgment and Decision Making Research in Accounting and Auditing*, edited by R. H. Ashton and A. H. Ashton. Cambridge, England: Cambridge University Press.

Von Winterfeldt, D., and W. Edwards. 1986. *Decision Analysis and Behavioral Research*. Cambridge, England: Cambridge University Press.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 804–824

The Effects of Error Frequency and Accounting Knowledge on Error Diagnosis in Analytical Review

Mark W. Nelson
Cornell University

SYNOPSIS AND INTRODUCTION: The performance of audit tasks has been modeled as a function of the auditor's ability, knowledge and experiences (Libby 1993). Thus, an important aspect of assigning audit tasks is identifying the levels of knowledge and types of experience an auditor must have to achieve a sufficiently high level of performance (Abdolmohammadi and Wright 1987). An objective of the current study is to determine whether auditors' knowledge of basic accounting principles and error frequencies improves over the course of their early careers so as to enhance performance of a common analytical procedure, ratio analysis.

Research in psychology suggests that two characteristics of a task (say, analytical procedures) could diminish the accuracy with which auditors learn error frequencies from experience and apply their knowledge to a task. First, auditors' memories of financial statement errors are encoded while they perform other information-processing activities. These competing task demands could use enough of an auditor's information-processing capacity to diminish both the accuracy with which memory traces of errors are encoded and the accuracy of their knowledge of error frequency (Naveh-Benjamin and Jonides 1986). Second, auditors must consider a variety of evidence when performing analytical procedures. In diagnostic tasks like ratio analysis, inordinate attention is given to evidence that is highly diagnostic of low-frequency events, causing an "inverse base rate effect" in which auditors consider such events as more likely (Medin and Edelson 1988). Assessing the extent to which either of these characteristics

The financial assistance of the Arthur Andersen & Co. Foundation, editorial assistance of Linda Pike, and programming assistance of Steven J. Benda are gratefully acknowledged. This research has benefitted from the comments of Andy Bailey, Sarah Bonner, Rob Bloomfield, Ernest Hicks, Dan Jensen, Karl Hackenbrack, Bob Libby, Daphne Main, Bill Notz, Tom Nygren, Eric Spires, Ken Trotman, four anonymous referees, and workshop participants at the University of Arizona, Cornell University, Indiana University, the University of Iowa, The Ohio State University, the University of Southern California, and the University of Wisconsin. Thanks also to Arthur Andersen & Co., Coopers & Lybrand, and Price Waterhouse for providing subjects for the experiment.

Submitted February 1991.
Accepted April 1993.

of the analytical procedures context prevents auditors from learning and applying error frequency knowledge is a second objective of this study.

To test hypotheses about these objectives, an experiment is conducted in which experienced auditors, accounting students, and non-accounting students learn the frequencies of financial statement errors through their experience in solving a series of problems using ratio analysis. Subjects are then tested for their accuracy in using frequency information by having them diagnose novel combinations of the same evidence. Other subjects perform similar tasks for an abstract medical diagnosis to provide a benchmark for comparison.

The results indicate that differences in accounting knowledge influenced the subjects' performance of ratio analysis, and that neither potential source of inaccurate learning of event frequency knowledge holds in this setting. That is, subjects learned frequencies in the presence of competing task demands, and the inverse-base rate effect was not observed. These results suggest that experienced, but not novice, auditors use both their superior knowledge of accounting and of error frequencies learned through experience. Another implication is that the performance of novice auditors may be improved by increasing their knowledge of basic accounting principles and error frequencies.

Key Words: *Analytical procedures; Frequency; Learning; Knowledge.*

Data Availability: *Data analyzed in this study may be obtained from the author on request.*

THE remainder of this article is organized as follows. Section I contains a review of relevant literature and a description of the hypotheses. The method by which these hypotheses are tested is discussed section II. Results are described in section III. A summary and general discussion appear in section IV.

I. Hypotheses and Related Literature

Accounting Knowledge and Performance of Analytical Procedures

The auditor uses his or her knowledge of accounting principles to determine those financial statement errors that are consistent with a pattern of evidence (Bedard and Biggs 1991). A poor knowledge of accounting principles can lead the auditor to investigate, and potentially accept, incorrect error diagnoses because of an incorrect integration of evidence.

In prior studies on the subject, Bedard and Biggs (1991) analyzed the verbal protocols of auditors (seniors and managers) performing a ratio analysis task and found that poor performance under some circumstances was caused by either the use of an incorrect definition of a ratio or the failure to integrate ratios into the correct pattern. They attributed these results to some subjects having difficulty accessing their knowledge of basic accounting principles, although differences in knowledge were not manipulated or measured directly as in Bonner and Lewis (1990), who correlated the performance of students and audit managers who completed the Bedard and Biggs (1991) task with their performance on a multiple choice test of accounting knowledge.

Bonner and Lewis found a significant correlation between performance on the two tasks for experienced auditors but not for students.

Libby and Tan (1993) reanalyzed the Bonner and Lewis (1990) data according to a model relating ability, experience, knowledge, and performance discussed in Libby (1993). This reanalysis demonstrated that it was the subjects' self-reported level of knowledge, not the knowledge proxy obtained from the multiple choice test administered by Bonner and Lewis (1990), that was significantly correlated with performance.

The current study complements these prior studies by taking a somewhat different approach: the amount of accounting knowledge is manipulated by using groups of subjects who differ in their experience and training. Also, a relatively structured task is used that does not require subjects to search for information or generate hypotheses. A relationship between differences in accounting knowledge and performance would be suggested by support for the following hypothesis:

H1: Performance in ratio analysis improves with knowledge of accounting principles.

Frequency Knowledge and Performance of Analytical Procedures

Prior Accounting Research. Using frequency information when performing analytical procedures can increase both audit effectiveness and efficiency. From an effectiveness standpoint, diagnostic failure in other settings has been shown to occur most often when the true explanation is omitted from a set of hypotheses (Elstein et al. 1978). This finding, coupled with the danger of accepting false but plausible explanations for evidence from analytical procedures (Kinney 1987), suggests that failure to include the most likely explanations in an initial set of hypotheses might result in accepting an incorrect hypothesis (Heiman 1990; Libby 1985). From an efficiency standpoint, failure to first investigate the most frequently occurring explanations would, on average, lead the auditor to expend unnecessary effort (Heiman 1990).[1]

This concern for audit efficiency and effectiveness motivated a study by Butt (1988) that examined learning of error frequency. She noted that auditors could acquire error frequency knowledge directly by experiencing financial statement errors on audits, or indirectly by experiencing errors in the form of error frequency statistics. Butt also noted that error frequency knowledge would be learned most accurately through direct experience if multiple trace theories of error frequency memory (Hintzman 1976; Hintzman et al. 1982) are true. She, therefore, compared the error frequency knowledge of auditors who were shown slides representing individual error causes (direct experience) with the error frequency knowledge of auditors who were shown slides that listed the number of occurrences of each error cause (indirect experience). Her results confirmed the hypothesis that auditors learned most accurately from direct experience.

Several studies have examined knowledge of error frequencies. Libby (1985) demonstrated that auditors have some knowledge of error-cause frequencies, for example, the frequency with which goods returned by customers are not recorded in the correct period, and that error causes that are perceived to occur more frequently are more likely to be generated as initial hypotheses. Libby and Frederick (1990) demonstrated that this knowledge of error-cause frequencies improves over broad levels of experience (accounting students, audit staff, and audit managers). Ashton (1991), however, found a relatively high knowledge of error-effect frequencies (e.g., the frequency with which inventory is misstated) in only the manufacturing and merchandising industries an, even in those industries, relatively accurate error frequency knowledge was evidenced for only the most frequently occurring error effects and causes.[2] Following the Hasher and Zacks' (1979, 1984) "automaticity hypothesis" and noting that auditors reported participation in relatively few audits per year, Ashton concluded that inaccurate frequency knowledge can be attributed to auditors not directly experiencing errors in proportion to their population frequencies.

---

[1] A similar point has been made in the cost variance investigation literature (Demski 1970; Jensen and McKeown 1977).

This study complements prior studies in two respects. First, as noted below, the evidence in psychology brings into question the automaticity hypothesis and suggests that some features of the auditing context might impede the learning of error frequencies from experience. The current study, therefore, tests the generality of the findings in Butt (1988) in a setting that operationalizes these features. Second, the relation between error frequency knowledge and hypothesis generation could be explained by auditors considering only the inherent risks associated with various transaction cycles when responding to both the hypothesis generation and error frequency estimation tasks, rather than considering error frequencies when generating hypotheses. Similarly, the improvement in error frequency knowledge with experience could be explained by the improvement in the auditor's knowledge of inherent risks over time. The current study manipulates error frequencies experimentally and thus provides experimental validation of the use of error frequency knowledge in analytical procedures.

The Influence of Features of the Auditing Context on Frequency Knowledge. Recent research in psychology has produced results questioning the automaticity hypothesis (for a review see Nelson 1993), suggesting other sources for inaccurate error frequency knowledge in auditing contexts. A possible explanation emerging from this research relates to the distinctiveness of memory traces of events that have been experienced. When attention is directed at some event (Hintzman 1976; Hintzman et al. 1982), a variety of factors may influence the encoding and retrieval of the trace of the event in memory, which, in turn, may influence the accuracy of error frequency information obtained from the trace itself (Begg et al. 1986). Two characteristics of the auditing environment that might influence the encoding of memory traces are the presence of competing task demands and the competition among different items of evidence for attention.

Competing Task Demands. As noted previously, the only prior error frequency learning research in accounting (Butt 1988) had subjects acquire error frequency information by viewing slides of error causes, with no concurrent information-processing activities taking place. In contrast, auditors performing analytical procedures must attend to both the causes of errors that actually occur and the evidence pointing to the presence of an error. They also must use their accounting knowledge to relate evidence and error causes when making diagnostic judgments.

---

[2] Ashton also detected no relationship between error-effect frequency knowledge and experience using several experience measures, including the broad levels of experience used by Libby and Frederick (1990). The reason for the discrepancy between Ashton's results that knowledge does not improve with experience for error effects and Libby and Frederick's results that knowledge does improve with experience for error causes is unclear.

Attending to evidence and engaging in effortful diagnostic activity could reduce the amount of attention and information-processing resources devoted to encoding traces of errors in memory. This, in turn, could reduce the degree to which the memory traces from which error frequency information is obtained can be discriminated, and thus reduce the accuracy with which error frequency information is later used in decision making. Prior studies in psychology support this view by finding decreased accuracy of error frequency knowledge when subjects observed events while also engaging in other effortful information-processing tasks, such as digit detection (Fisk and Schneider 1984), counting (Naveh-Benjamin and Jonides 1986), and addition (Maki and Ostby 1987). To the extent that an auditor's attention to evidence and diagnosis activity diminishes the accuracy with which errors that had been encountered are encoded in memory, a similar decrease in the accuracy of error frequency knowledge could occur. Thus, it is unclear that error frequency will be learned accurately enough in the presence of competing task demands to improve performance of subsequent analytical procedures. The hypothesis to be tested is as follows:

H2: Error frequencies are learned when experienced in the course of making diagnostic judgments.

Competition Among Pieces of Evidence. Other research suggests that, because subjects in psychology must distribute their attention among several items of evidence, they might, under some circumstances, inversely apply the error frequency information to which they have been exposed. This inverse base rate effect was originally discovered by Medin and Edelson (1988) in tests of the influence of error frequency information on categorization judgments. Their experiments were divided into two phases. During the learning phase, subjects were presented with a series of learning trials designed to teach them (1) disease frequencies and (2) the symptoms diagnostic of particular diseases. On a given trial, a subject would be presented with a pair of symptoms, asked to diagnose which of several imaginary diseases was suggested by the symptoms, and be given feedback as to the correct answer. After several repetitions, for example, a subject might learn that disease 1 (althrax) was indicated by symptoms *A* (bloody nose) and *B* (stomach cramps), disease 2 by *A* and *C*, disease 3 by *D* and *E*, and disease 4 by *D* and *F*. In this example, symptoms *B*, *C*, *E*, and *F* are highly diagnostic, because each is uniquely associated with its respective disease, while symptoms *A* and *D* are less diagnostic, each being associated with two diseases. The symptoms and diseases used by Medin and Edelson (1988) are shown in table 1. Disease frequency was manipulated by having one of the symptom/disease combinations appear three times more often than the other (e.g., *A* and *B*, indicating disease 1, appeared three times more often than *A* and *C*, indicating disease 2). Three sets of these symptom/disease combinations were presented in the design shown in table 2.

After subjects had reached a high level of performance in the learning phase, they entered the transfer phase of the experiment. In the transfer phase, 33 new combinations of the symptoms that appeared during the learning phase were presented for diagnosis, as shown in table 3. These new combinations were designed to force subjects to apply what they had learned in the learning phase. The results revealed that, when presented with single highly diagnostic symptoms (e.g., *B* or *C*), subjects

Table 1
Symptoms and Imaginary Diseases[a] Used in the Medical Diagnosis Context

First Cue Combination

A. Bloody nose
B. Stomach cramps
⇒ 1. althrax

A. Bloody nose
C. Irritated eyes
⇒ 2. buragamo

Second Cue Combination

D. Discolored gums
E. Swollen tongue
⇒ 3. burlosis

D. Discolored gum
F. Itchy throat
⇒ 4. coralgia

Third Cue Combination

G. Sweaty palms
H. Aching back
⇒ 5. midosis

G. Sweaty palms
I. Dry skin
⇒ 6. terrigitis

[a] Adapted from Medin and Edelson (1988).

responded with the correct diagnosis, demonstrating that they had learned the symptom that was highly diagnostic of each disease. When presented with single less-diagnostic evidence tests (e.g., *A* or *D*) or combined tests (e.g., *A*, *B*, and *C* together), subjects most often selected the most frequent disease, demonstrating that they had learned disease frequency information. But when presented with tests that paired highly diagnostic evidence (e.g., *B* and *C*) or tests that paired highly diagnostic evidence with novel, less-diagnostic evidence (e.g., *B* and *D*), subjects diagnosed the *less* frequently occurring disease more often than the *most* frequently occurring disease. Thus, after demonstrating that they had acquired disease frequency knowledge, subjects violated that knowledge in response to certain combinations of items of evidence. In a situation in which disease frequency information and diagnostic evidence information were learned jointly and applied jointly in diagnosis, disease frequency information was applied incorrectly in some circumstances.

This result, labeled "the inverse base rate effect," could not be explained by existing categorization models, which led Medin and Edelson (1988) to modify Medin and Schaffer's (1978) context model to include assumptions about learning that would account for the inverse base rate effect. The modified context model pictures a decision maker as reacting to feedback that he or she has made a misdiagnosis by allocating more attention in the future to those items of highly diagnostic evidence that sup-

Table 2
Relative Frequencies of Ratio/Error or Symptom/Disease Pairings[a]

Relative Frequency in Learning Stage	*Symptom/Ratio*	*Underlying Disease/Error*
3	A,B[b]	1
1	A,C	2
3	D,E	3
1	D,F	4
3	G,H	5
1	G,I	6

[a] Adapted from table 1 in Medin and Edelson (1988, 71).

[b] Letters match those of the ratio differences in table 4 for the analytical procedures context and those of symptoms in table 1 for the medical diagnosis context.

Table 3
Relation of Problem Types in Transfer Phase to Hypotheses[a]

Problem Type	*Description*[b]	*Hypothesis*
1	Single Highly Diagnostic Evidence Tests: B, C, E, F, H, I	Manipulation Check
2	Single Less-Diagnostic Evidence Tests: A, D, G	H2 (error frequency knowledge)
3	Combined Highly and Less-Diagnostic Evidence Tests: ABC, DEF, GHI	H2 (error frequency knowledge)
4	Conflicting Highly Diagnostic Evidence Tests: BC, BF, BI, EC, EF, EI, HC, HF, HI	H3 (inverse base rate effect)
5	Conflicting Highly and Less-Diagnostic Evidence Tests: AE, AF, AH, AI, DB, DC, DH, DI, GB, GC, GE, GF	H3 (inverse base rate effect)

[a] Adapted from table 1 in Medin and Edelson (1988, 71). Hypothesis H1 is tested with performance data from the learning phase.

[b] Letters match those of the ratio differences in table 4 for the analytical procedures context and those of symptoms in table 1 for the medical diagnosis context.

ported the correct diagnosis.[3] The resulting increased salience of this highly diagnostic evidence would make it more persuasive in subsequent categorization judgments. It was also assumed that more errors would be made when diagnosing low-frequency events than high-frequency events. These two assumptions imply that highly diagnostic evidence that supports a low-frequency event is more salient (gets more attention) than would be similar evidence supporting a high-frequency event. Thus, when the two types of evidence conflict, the low-frequency event would be judged the more likely. That is, over time a decision maker learns that the high-frequency event is indicated by the entire set of evidence, and the low-frequency event is the exception indicated by its particular highly diagnostic evidence (Medin and Bettger 1991).

---

[3] See Gluck and Bower (1988a, 1988b), Estes et al. (1989), and Medin and Bettger (1991) for other results suggesting that allocation of attention among evidence values can explain the neglect of error frequency information in categorization.

Several characteristics of the analytical procedures context parallel those in the medical diagnosis context used by Medin and Edelson (1988) and thus might allow the process described in the modified context model (and the resulting inaccurate use of event frequency information in response to some evidence) to take place. First, financial statement errors produce evidence, such as unexpected ratio values, just as diseases produce symptoms. Second, auditors observe unexpected ratio values and try to determine what underlying errors might be present in the financial statements, just as doctors diagnose diseases based on observed symptoms. Third, auditors access their knowledge to help them perform this diagnosis, and their knowledge is, at least in part, a product of their past audit experiences. Thus, auditors performing analytical procedures are likely to consider previous judgments when presented with similar sets of evidence. Auditors are also expected to use that knowledge to help guide the current diagnostic judgment. This is similar to the experience-based reasoning necessary in the medical diagnosis context. Fourth, after performing analytical procedures during audit planning, an auditor directs audit effort toward determining whether the auditor's error diagnoses are correct. Some procedures, such as inquiry, might provide immediate feedback, while others might only provide feedback after subsequent audit tests have been completed. That is, feedback is eventually made available. Fifth, as in the medical diagnosis context, auditors constantly face new configurations of evidence, especially because of non-error fluctuations in financial ratios. These similarities between the analytical procedures and medical diagnosis contexts suggest that the process described in the modified context model might take place in the analytical procedures context. This implies the following hypothesis:

H3: The inverse base rate effect occurs in the analytical procedures context.

II. Method

Experimental Design and Independent Variables

A 3×2 between-subjects design was used in the experiment. Accounting Knowledge (high, low, none) allowed the manipulation of accounting knowledge and, thus, avoided the construct-validity problems with knowledge measures observed in prior studies. Context (analytical procedures, medical diagnosis) varied subjects' ability to use their accounting knowledge in diagnosis and performed two useful functions. First, it replicated prior psychology research in the medical diagnosis context, so that any lack of inverse base rate effect in the auditing context could be reasonably attributed to a difference in contexts. Second, the medical diagnosis context provided a benchmark to control for differences in ability or motivation across groups in interpreting knowledge effects. In this approach, discussed by Bonner (1990) and Frederick and Libby (1986), the basic characteristics of the judgment task are held constant between contexts to avoid confounding the effects of knowledge differences with other task characteristics. The knowledge of interest (in this case, accounting knowledge) is useful in one context but not the other. Differences in subjects' ability or motivation are suggested by a main (or simple) effect for subject type, differences attributable to context are suggested by a main (or simple) effect for context, and a context/subject-type interaction demonstrates the effect of knowledge differences.

Subjects

Seventy-four auditors from three firms participated as subjects with a "high" level of accounting knowledge.[4] The firms were requested to distribute data collection packets to auditors with more than two years experience, the level at which auditors typically perform analytical procedures. Auditors' experience had a mean of 47.03 months (standard deviation 22.29 months), with a low of 9 months and a high of 130 months.[5]

One hundred and twelve intermediate accounting students participated as subjects with a "low" level of accounting knowledge. The accounting students were recruited from their first intermediate accounting course, of which they had completed approximately 75 percent. These students were chosen because their accounting training should have exposed them to the GAAP necessary to understand the analytical procedures context. Fifty-seven other intermediate-course students with similar backgrounds provided data at a later date for a manipulation check.

Eighty-seven nonaccounting students participated as subjects with no accounting knowledge. These students were enrolled in nonbusiness undergraduate majors. Thirteen of these subjects indicated that they had either prior accounting coursework or accounting experience and so were assigned to the medical diagnosis context. All nonaccounting subjects assigned to the analytical procedures context had no prior accounting coursework or experience.

Procedure

Data were collected using an interactive computer program run on microcomputers. The data collection program maintained control over the order in which participants completed the experimental tasks, required participants to complete the materials in one sitting, did not allow incomplete responses, and prevented any changing of responses once participants had answered a question. Auditors participating in the experiment completed the data collection materials at their offices, while students completed the materials during scheduled sessions with the author.

The data collection program proceeded as follows. After inputting an identification code, subjects first viewed a short introduction and instructions for the learning phase of the experiment. They then completed a number of learning trials. In each learning trial, a subject viewed a set of unexpected ratio values (or symptoms, if the subject was assigned to the medical diagnosis context) and selected from a list of six errors (or diseases) the one that he or she believed was most likely to have caused the unexpected ratio values (or symptoms). The subject then received feedback as to whether his or her response was correct. The feedback always repeated the information on which

---

[4] Fifty-one subjects were obtained from six offices of the first firm, 20 subjects were obtained from one office of the second firm, and three subjects were obtained from one office of the third firm. Because (1) it was cognitive processes that were being examined, (2) no firm effects had been hypothesized, and (3) there were differences in numbers of subjects from each firm, no analysis of firm effects were conducted.

[5] Four auditors had less than 24 months of audit experience. No results changed with these subjects omitted, and the results reported include them in the analyses.

Table 4
Combinations of Ratio Differences Suggesting Financial Statement Errors[a]

	Expected[b]	Actual[c]
First Cue Combination		
A. Return on Sales	0.0432	0.0466
B. Net Fixed Assets to Gross Fixed Assets	0.6310	0.6500
⇒ 1. Depreciation Expense for 1988 understated		
A. Return on Sales	0.0432	0.0466
C. Other Liabilities to Total Liabilities and Equity	0.0393	0.0322
⇒ 2. Pension Expense for 1988 understated		
Second Cue Combination		
D. Quick Ratio	1.166	1.349
E. Average Collection Period	54.454	67.308
⇒ 3. Net Accounts Receivable at 12/31/88 overstated		
D. Quick Ratio	1.166	1.349
F. Times Interest Earned	2.943	5.210
⇒ 4. Financial Expense for 1988 understated		
Third Cue Combination		
G. Current Ratio	2.704	3.319
H. Inventory Turnover	2.562	1.808
⇒ 5. Cost of Goods Sold for 1988 understated		
G. Current Ratio	2.704	3.319
I. Long-Term Liabilities to Total Assets	0.190	0.240
⇒ 6. Current Maturities of Debt at 12/31/88 understated		

[a] In each case the item of less-diagnostic evidence is presented first, then the item of highly diagnostic evidence, and then the implied financial-statement error.

[b] "Expected" values are based on the company's 1987 audited financial statements, but also include the effects of known changes in circumstances and other relevant information such as budgets and forecasts.

[c] "Actual" values are calculated from the company's unaudited 1988 financial statements.

the subject based his or her response, state the correct response and repeated the subject's response if it was incorrect.

There were six different ratio analysis (or medical diagnosis) problems presented in the learning phase. The relative number of times that a subject was presented with each problem was varied between subjects to manipulate subjects' error or diagnosis frequency knowledge. For example, for one half of the subjects, the "Depreciation Expense understated" problem in table 4 occurred three times more often during the learning phase than did the "Pension Expense understated" problem, while the reverse was affected for the other half of subjects.[6] Twelve learning trials were necessary to convey 3:1 relative frequencies for three combinations of high- and low-frequency errors (or diseases). The structure of the learning phase parallels that used by Medin and Edelson (1988). To control for order effects, the order of the learning trials was randomized within each block of twelve trials, and the order in which the individual ratios (or symptoms) appeared and in which the error (or disease) alternatives from which a subject could select a response appeared were randomized within each trial. The learning phase ended for a subject when one of two events had occurred: either (1) the subject was correct on every diagnosis for 24 consecutive trials, or (2) the subject completed 16 blocks of 12 trials.

---

[6] Manipulating error frequencies provided a control for the real-world error frequency knowledge of experienced auditors. If real-world error frequencies dominate experimental error frequencies, auditors would demonstrate no knowledge of experimental error frequencies.

After exiting from the learning phase, subjects viewed another set of instructions and entered the transfer phase. In the transfer phase, subjects completed 33 additional problems that presented them with new combinations of the ratio values (or symptoms) that subjects encountered during the learning phase. The problems were designed to determine whether subjects would use the error (or disease) frequency information they acquired during the learning phase. To control for order effects, the order in which unexpected ratio values (or symptoms) appeared and the order in which the six alternative responses appeared were randomized within each trial, and the order in which the 33 problems appeared in the transfer phase was randomized for each subject. At the completion of the transfer phase, subjects exited from the data collection program and completed a debriefing questionnaire. Nonaccounting subjects assigned to the medical diagnosis context also completed a short accounting test to provide data for a manipulation check.

Experimental Materials

Data collection materials for subjects assigned to the analytical procedures context will be described first, then materials relevant to the medical diagnosis context.

All subjects received a packet containing a cover letter, instructions on how to activate the data collection program, a diskette containing the data collection program, a packet of reference materials, and a debriefing questionnaire. The cover letter thanked the participants for their effort and gave detailed instructions, describing the procedures necessary to activate the data collection program. Auditors' letters also (1) stressed that their firm had determined that the project was of practical importance, (2) informed them that they had to complete the experiment when they had a two-hour block of uninterrupted time available, and (3) provided a toll-free telephone number at which to call the author if they had trouble activating the data collection program. Students were told that they would receive bonus points from their instructors and a five dollars cash payment for their participation.

Instructions for the learning phase informed subjects in the analytical procedures context that they were engaged in a training exercise designed to help them use ratio analysis to identify possible material errors in the financial statements of firms in the widget manufacturing industry and that the (imaginary) widget industry was selected so that all participants would have the same amount of knowledge about the companies used in the experiment. Frequencies or evidence diagnosticities were not mentioned. The audited 1987 Income Statement and audited 12/31/87 Balance Sheet of a typical widget manufacturer, a glossary of financial ratio definitions, and the instructions were provided to subjects as handouts for use throughout the experiment.

The combinations of unexpected ratio values and financial statement errors that were used in the learning phase are shown in table 4. Each of the six errors either overstated income or liquidity. Unexpected ratio values were conveyed by presenting subjects with both an expected and actual value for each ratio presented. The expected 1988 ratio values were computed from the same 1987 financial statements supplied to subjects, after considering obvious changes between years, such as reducing "Long-Term Liabilities" for payment of 1987 "Current Maturities of Long-Term Debt" and increasing "Accumulated Depreciation." Actual ratio values were computed in the same way as expected ratio values, but after seeding the relevant material errors. Materiality guidelines of 5 percent of Operating Income and 5 percent of Total Assets (Arens and Loebbecke 1991) were used to ensure that the ratio differences were of the correct magnitudes. Subjects were not informed of these guidelines but were informed that they were dealing with a material error. In some cases the seeded error was greater than the materiality guideline, to ensure that the single items of less-diagnostic evidence for a pair of high/low frequency errors had the same value for each of the errors.[7]

Instructions for the transfer phase informed subjects that they had completed the learning phase, that they were now entering the transfer phase, and that the objective of the transfer phase was to find out how well they were able to use the experience gained in the learning phase. Subjects were also informed that all previous instructions were still in effect, that what they had seen in the learning phase constituted a representative sample of companies (or patients), and that the experience they gained in the learning phase was therefore relevant to the companies (or patients) they would see in the transfer phase.

Following Medin and Edelson (1988), subjects were given a total of 33 problems to complete in the transfer phase. Each problem was a new combination of the same items of evidence that subjects encountered during the learning phase. Items of evidence that were always presented in pairs during the learning phase were presented individually, in new pairs, or even in triplets in the transfer phase. The problems were designed to require subjects to use, and thus reveal, the error (or disease) frequency knowledge that they had acquired in the learning phase. The problems fell into five categories, as shown in table 3 above, and are described under "Dependent Variables" below.

Before any data were collected, all auditing materials were examined for clarity and reasonableness by a former senior technical partner of a Big-6 public accounting firm. After extensive discussion and revision, both the analytical procedures and medical diagnosis data collection materials were pilot tested with both auditors and students. Based on feedback from pilot testing, the wordings of instructions, learning tasks, transfer tasks and debriefing questions were changed to reduce apparent ambiguities. These changes were reviewed and approved by the partner who reviewed the original materials.

The data collection materials for the medical diagnosis context were similar to those in the analytical procedures context, except that symptom combinations replaced ratio difference combinations and imaginary diseases replaced financial-statement errors. Table 1 above shows the combinations of symptoms and imaginary diseases used in the experiment.

---

[7] Computations of the expected and actual ratio values are available from the author upon request.

A short accounting test was administered to provide a manipulation check that (1) the nonaccounting students had no accounting knowledge, and (2) the intermediate accounting students had accounting knowledge. The test contained eight multiple choice questions about some of the basic accounting concepts and financial ratio definitions that would be included in the GAAP knowledge relevant to the analytical procedures context. Each question had four alternative answers; a score of 2 out of 8 constitutes performance at a chance level.

Dependent Variables

Performance data from the learning phase were used to test hypothesis H1, and data from the transfer phase were used to test hypotheses H2 and H3. In the transfer phase, subjects responded to five types of problems (table 3): one type provided data for a manipulation check, two types provided data for tests of frequency learning, and two provided data for tests of the inverse base rate effect.

Dependent Variable for Manipulation Check. To determine whether, after exiting from the learning phase, subjects knew the relationships between errors (or diseases) and unexpected ratio values (or symptoms), they were presented with six single items of highly diagnostic evidence (e.g., *B* in table 3 above; *B* would correspond to "stomach cramps" in table 1 for the medical diagnosis subjects, and "Net Fixed Assets to Gross Fixed Assets higher than expected" in table 4 for the analytical procedures subjects). A diagnosis in response to one of these predictors was judged correct if it was the same event paired with the predictor during learning (e.g., as shown in table 1, if presented with "stomach cramps," the correct response would be "althrax"; as shown in table 4, if presented with "Net Fixed Assets to Gross Fixed Assets higher than expected," the correct response would be "Depreciation Expense for 1988 understated"). Correct diagnoses in response to these problems would indicate that subjects understood the errors (or diseases) that were suggested by particular unexpected ratio values (or symptoms). Thus, the dependent variable used was the proportion of times over the six problems that a subject answered with the correct response.

Dependent Variables for Tests of Accounting Knowledge (H1). The numbers of misdiagnoses that subjects made in the learnings phase were used to determine whether knowledge differences influence performance in either context.

Dependent Variables for Tests of Event Frequency Knowledge (H2). Two problems examined whether subjects learned the event frequency information that was presented in the learning phase. Subjects were presented with single items of less-diagnostic evidence (e.g., *A* in table 3 above) and three combinations of two items of conflicting highly diagnostic evidence and their common item of less-diagnostic evidence (e.g., *A*, *B*, and *C* in table 3 above). For each of the two problems, the dependent variable used was the difference between the proportion of times subjects diagnosed the high- and low-frequency events associated with the predictors during learning.

Dependent Variables for Tests of the Inverse Base Rate Effect (H3). Two types of problems were used to examine inverse base rate effects. In each case, the problem was designed to determine whether evidence that was highly diagnostic of low-frequency events was more persuasive than evidence that was highly diagnostic of high-frequency events. The first type of problem presented subjects with nine pairs of conflicting items of highly diagnostic evidence (e.g., *B* and *C* in table 3 above). The difference between the proportions of times that subjects diagnosed the high- and low-frequency events associated with the predictors during learning was used as the dependent variable. The second type of problem presented subjects with pairs of conflicting items of highly diagnostic and less-diagnostic evidence that had not been paired together during the learning phase. Six of the twelve problems presented a highly diagnostic predictor of a low-frequency event (e.g., *F* of *A* and *F* in table 3 above is shown in table 2 as predicting a low-frequency event), and the other six presented a highly diagnostic predictor of a high-frequency event (e.g., *E* of *A* and *E* in table 3 above is shown in table 2 as predicting a high-frequency event). The inverse base rate effect would be indicated if there was a negative difference between the proportion of times a subject selected the event associated with the highly diagnostic evidence when the event was of high frequency and that proportion when the event was of low frequency.

III. Results

Manipulation Checks

Knowledge Differences. A short accounting test was administered to the nonaccounting students participating in the medical diagnosis context,[8] as well as to a group of 57 intermediate accounting students with the same prior accounting coursework as the intermediate students who participated in the analytical procedures or medical diagnosis contexts.[9] Results revealed a mean score of 5.63 for the intermediate accounting students and 2.125 for the nonaccounting students; two-sample *t*-tests reveal the scores to be significantly different ($t = 11.79$, $p < 0.0001$). The performance of the nonaccounting students was insignificantly different from a chance score of 2 ($t = 0.54$, $p > 0.5$).[10] Thus, as expected, students in intermediate accounting possessed more accounting knowledge than the nonaccounting students and that the nonaccounting students performed at a level that suggests no accounting knowledge.

Knowledge of Relationships Between Errors (or Diseases) and Ratios (or Symptoms). One type of problem administered during the transfer phase presented subjects with a single item of highly diagnostic evidence. Results for this problem type are shown in table 5. In all cases, subjects selected the response suggested by the evidence over 75 percent of the time, which was significantly greater than chance performance (a proportion of 1/6) for all subject groups. An ANOVA on the arcsine-transformed[11] proportion of correct responses revealed a significant Context by

---

[8] The test was not given to nonaccounting students assigned to the analytical review context because of concern that it might either (1) contaminate subjects' lack of accounting knowledge if given as a pretest, or (2) be contaminated by the learning phase if given as a posttest.

[9] These data cannot be used as knowledge proxies in the manner of Bonner and Lewis (1990), because (except for the nonaccounting students) the same subjects did not take the accounting knowledge test and participate in the learning and transfer phases.

[10] Thirteen of the nonaccounting subjects who took the test indicated that they had either prior accounting coursework or experience, so this is a conservative comparison. With these 13 subjects deleted, the mean score of the nonaccounting subjects drops to 1.67.

[11] The specific variance stabilizing transformation recommended for proportions is

$$h(t) = (2/a)[\sin^{-1}(t^{.5})]$$

(Bickel and Doksum 1977); "a" = pi is approximated by 3.141593.

Table 5
Test for Knowledge of Relationships Between Errors (or Diseases) and Ratios (or Symptom)[a]

Subject/Context[b]	n	*Proportion*[c]	t[d]	p-*value*
N, MD	41	0.882	28.07	<0.01
S, MD	56	0.860	28.53	<0.01
A, MD	36	0.856	31.14	<0.01
N, AP	46	0.772	16.98	<0.01
S, AP	56	0.809	25.42	<0.01
A, AP	38	0.906	38.14	<0.01

[a] Subjects presented with a single item of highly diagnostic evidence.

[b] N=nonaccounting student, S=accounting student, A=auditor, MD=abstract medical diagnosis context, and AP=analytical procedures context.

[c] Mean proportion of time subjects chose the response associated with the evidence.

[d] *t*-test comparing proportion to chance level of performance (1/6).

Accounting Knowledge interaction ($F = 4.67$, $p < 0.011$). Further analysis revealed a significant simple effect for Accounting Knowledge in the analytical procedures context ($F=4.93$, $p<0.01$) but not in the medical diagnosis context ($F=0.67$, $p>0.5$), with Tukey studentized range tests indicating that auditors performed significantly better than both intermediate and nonaccounting students ($p < 0.05$).

Tests for Effect of Differences in Accounting Knowledge

To test hypothesis H1, the learning performance of the three groups of subjects was examined for a knowledge effect. Although the ratio analysis task was structured so as to prevent subjects' abilities from affecting their performance, using the medical diagnosis context as a benchmark provided an added control.

Because each pair of ratio differences presented during the learning phase suggested only one of the six alternative responses, a correct diagnosis was indicated when a subject responded with the alternative suggested by the ratio differences, and a misdiagnosis was indicated when a subject responded with one of the other five alternatives. The mean values of the number of misdiagnoses made by each subject during the learning phase are displayed in table 6 by context and experience.[12] An ANOVA was performed with these values as the dependent variable and Accounting Knowledge and Context as independent variables. Results revealed a significant Accounting Knowledge by Context interaction ($F = 5.87$, $p < 0.01$). Analysis of simple effects revealed that Accounting Knowledge was not significant in the medical diagnosis context ($F = 1.21$, $p > 0.30$) but was significant in the auditing context

---

[12] Eleven subjects did not exit from the learning phase before completing 16 blocks of 12 trials. These subjects automatically exited from the learning phase after completing 192 trials. Eight of the subjects (four nonaccounting, four intermediate accounting) were assigned to the analytical procedures context, and three of the subjects (two nonaccounting, one intermediate accounting) were assigned to the medical diagnosis context. Also, seven subjects (six auditors assigned to the analytical procedures context, one intermediate accounting student assigned to the medical diagnosis context) made no misdiagnoses during the learning phase. Results did not change with either of these groups of subjects omitted from the analysis, and they are included in all of the analyses reported.

Table 6
Number of Errors Made During Learning Phase

Subject/Context[a]	*Number of Errors*
N, MD	15.90
S, MD	13.56
A, MD	9.89
N, AP	30.59
S, AP	19.88
A, AP	4.29

[a] N=nonaccounting student, S=accounting student, A=auditor, MD=abstract medical diagnosis context, and AP=analytical procedures context.

($F = 17.60$, $p < 0.0001$). Tukey studentized range tests showed no significant differences between levels of Accounting Knowledge in the medical diagnosis context, but significant differences between each level of Accounting Knowledge in the analytical procedures context ($p < 0.05$). These results support hypothesis H1.

Tests for Application of Frequency Knowledge

Two more types of problems administered during the transfer phase were designed to examine the accuracy of subjects' knowledge of the event frequencies to which they had been exposed in the learning phase of the experiment. One problem presented subjects with a single item of less-diagnostic evidence, while the other presented subjects with a combination of two conflicting items of highly diagnostic evidence and an item of less-diagnostic evidence that had been paired with both of the items of highly diagnostic evidence during learning. For each problem type, both a high-frequency event and a low-frequency event could be supported by the combination of evidence.

The proportion of times that subjects selected the high- and low-frequency responses is listed in table 7. One-sided t-tests were performed to determine whether the difference between high- and low-frequency responses was greater than zero. For the single less-diagnostic evidence problem, this difference was significantly above zero for all subjects except the auditors who were assigned to the analytical procedures context. For the combined highly- and less-diagnostic evidence problem, this difference was significantly above zero for all subjects. These results support hypothesis H2.

Tests for Inverse Base Rate Effects in the Analytical Procedures Context

Two problem types are relevant to determining whether auditors are vulnerable to the inverse base rate effect in the analytical procedures context: conflicting highly-diagnostic evidence problems and conflicting highly- and less-diagnostic evidence problems.

Table 7
Tests for Event Frequency Knowledge

Subject/ Context[a]	n	Proportion of Times Subjects Chose: High Freq. Response	Proportion of Times Subjects Chose: Low Freq. Response	Diff.[b]	t	p-value
1. Single Less-Diagnostic Evidence Tests[c]						
N, MD	41	0.642	0.211	0.431	5.62	<0.01
S, MD	56	0.637	0.161	0.476	8.04	<0.01
A, MD	36	0.565	0.315	0.250	2.49	<0.01
N, AP	46	0.377	0.174	0.203	3.07	<0.01
S, AP	56	0.333	0.161	0.172	3.61	<0.01
A, AP	38	0.303	0.263	0.040	0.60	>0.25
2. Combined Highly and Less-Diagnostic Evidence Tests[d]						
N, MD	41	0.561	0.398	0.163	2.02	<0.06
S, MD	56	0.643	0.292	0.351	4.91	<0.01
A, MD	36	0.667	0.305	0.362	4.38	<0.01
N, AP	46	0.515	0.355	0.160	2.00	<0.06
S, AP	56	0.554	0.315	0.239	3.19	<0.01
A, AP	38	0.548	0.417	0.131	1.63	<0.06

[a] N=nonaccounting student, S=accounting student, A=auditor, MD=abstract medical diagnosis context, and AP=analytical procedures context.

[b] Mean of the difference for each subject between the proportion of times the subject chose the high-frequency response and the proportion of times the subject chose the low-frequency response.

[c] Subjects presented with a single item of less-diagnostic evidence.

[d] Subjects presented with three items of evidence: one that is highly diagnostic of a high-frequency event, one that is highly diagnostic of a low-frequency event, and one that is less diagnostic of either event.

When responding to conflicting highly-diagnostic evidence problems, the inverse base rate effect would be demonstrated if a subject chose the low-frequency event more often than the high-frequency event. Therefore, one-sided *t*-tests were used to test the hypothesis that the difference between the proportion of times a subject selected the high-frequency event and the proportion of times a subject selected the low-frequency event was less than zero. As shown in table 8, subjects did not exhibit significant inverse base rate effects in the analytical procedures context, but did exhibit the inverse base rate effect in the medical diagnosis context.

When responding to conflicting highly- and less-diagnostic evidence problems, the inverse base rate effect would be suggested if the proportion of times a subject selected the event supported by the highly diagnostic evidence was greater when the event was of low frequency than when it was of high frequency. Therefore, a difference score was computed by subtracting the proportion of times a subject selected the response associated with the highly diagnostic evidence when the evidence was of a low-frequency event from that proportion when the evidence was of a high-frequency event, and one-sided *t*-tests were used to test the hypothesis that this difference was less than zero. As shown in table 8, subjects did not exhibit significant inverse base rate effects in the analytical procedures context, but did exhibit inverse base rate

Table 8

Tests for the Inverse Base Rate Effect

1. Conflicting Highly Diagnostic Evidence Tests[a]

		Proportion of Times Subjects Chose				
Subject/ Context[b]	n	*High Freq. Response*	*Low Freq. Response*	*Diff.*[c]	t	p-*Value*
N, MD	41	0.393	0.538	−0.145	−2.28	<0.02
S, MD	56	0.413	0.529	−0.116	−1.69	<0.05
A, MD	36	0.401	0.551	−0.150	−1.52	<0.10
N, AP	46	0.443	0.442	0.001	0.01	>0.5
S, AP	56	0.448	0.404	0.044	0.86	>0.5
A, AP	38	0.483	0.439	0.044	0.50	>0.5

2. Combined Highly and Less-Diagnostic Evidence Tests[d]

		Proportion of Times Subjects Chose Response Associated with Highly Diagnostic Evidence When Response Was of				
Subject/ Context[b]	n	*High Freq. Response*	*Low Freq. Response*	*Diff.*[e]	t	p-*Value*
N, MD	41	0.622	0.675	−0.053	−1.37	<0.10
S, MD	56	0.554	0.661	−0.107	−2.30	<0.05
A, MD	36	0.551	0.625	−0.074	−1.39	<0.10
N, AP	46	0.630	0.583	0.047	1.22	>0.5
S, AP	56	0.670	0.625	0.045	1.08	>0.5
A, AP	38	0.706	0.726	−0.020	−0.40	>0.30

[a] Subjects presented with an item of evidence that is highly diagnostic of a high-frequency event and an item of evidence that is highly diagnostic of a low-frequency event.

[b] N=nonaccounting student, S=accounting student, A=auditor, MD=abstract medical diagnosis context, and AP=analytical procedures context.

[c] Mean of the difference for each subject between the proportion of times the subject chose the high-frequency response and the proportion of times the subject chose the low-frequency response.

[d] Subjects presented with two items of evidence: one that is highly diagnostic of *either* a low or high frequency event, and one that is less diagnostic of a different event.

[e] Mean of the difference for each subject between the proportion of times and the subject chose the response associated with highly diagnostic evidence when the response was of high frequency and that proportion when the response was of low frequency.

effects in the abstract medical diagnosis context. These results do not support hypothesis H3.[13]

An ANOVA performed for each of the two types of problems clarifies the difference in results across contexts. In each case the difference measure previously described served as the dependent variable and Context and Accounting Knowledge served as independent variables. Results for the two problem types were quite similar. In predicting difference scores for conflicting, highly-diagnostic evidence problems, Context was significant ($F = 7.03$, $p < 0.01$), Accounting Knowledge was not significant ($F = 0.12$, $p > 0.85$), and the Context by Accounting Knowledge interaction was not significant ($F = 0.05$, $p > 0.95$). In predicting difference scores for conflicting highly- and less-diagnostic evidence problems, Context was significant ($F = 7.58$, $p < 0.01$), Accounting Knowledge was not significant ($F = 0.45$, $p > 0.6$), and the Context by Accounting Knowledge interaction was not significant ($F = 0.59$, $p > 0.55$).

---

[13] While the means in table 8 indicate that there is no inverse base rate effect in the analytical procedures context, they also show an indifference between the high- and low-frequency responses, which would not be suggested if subjects were actually using the frequency information. This may have occurred

IV. Discussion

One finding of this study is that auditors' knowledge of very basic accounting principles improves with work experience from the level attained in college coursework, and thus enhances their performance of ratio analysis. This result supports the suggestion of Bedard and Biggs (1991) and the findings of Bonner and Lewis (1990) that differences in basic accounting knowledge explain a portion of the variance in the performance of analytical procedures. Future research should consider the particular aspects of accounting knowledge that improve over an auditor's career, as well as the costs and benefits of complementing the accounting knowledge of novice auditors with additional training. If performance gains resulting from increased training are significant enough, analytical procedures might be assigned to less-experienced personnel with a minimal decrease in audit effectiveness and efficiency.

A second finding of this study is that auditors can acquire event frequency knowledge from their experience in performing analytical procedures and that they apply this event frequency knowledge in future analytical procedure judgments.[14] This result indicates that the existence of competing task demands did not prevent event frequency learning, which supports the generality of the findings of Butt (1988) in a more realistic learning setting, and also provides experimental validation of the correlational results of Libby (1985) and Libby and Frederick (1990). It also implies that auditors' knowledge of error frequencies and evidence diagnosticities might be enhanced through the use of simulated audit experience similar to the computer program used for this experiment. As noted by Ashton (1991), auditors encounter relatively few instances of errors in actual audit practice, so the error frequency knowledge they acquire from experience may not reflect population frequencies. Simulated audit experience might be a cost-effective method of increasing the set of experiences on which the auditor can draw.

A third finding of this study is a lack of inverse base rate effect for all subjects in the analytical procedures setting, but a strong inverse base rate effect for all subjects

---

because the novel combinations of accounting evidence were meaningful, and therefore may have suggested to subjects an alternative not included among the list of responses. This effect is discussed further in the discussion section.

[14] The only case in which a subject group did not demonstrate significant event frequency learning was when auditors assigned to the analytical procedures context responded to single less-diagnostic evidence problems, and this probably occurred because the single items of less-diagnostic evidence were sufficiently vague to be consistent with virtually all of the financial statement errors presented as alternative responses. For example, the higher-than-expected Return on Sales is consistent with all error alternatives in table 4 except for "Current Maturities of Debt at 12/31/88 understated."

in the abstract medical diagnosis setting. Although the learning phase of the experimental task differed from audit practice by providing a large number of exposures to error and rapid, accurate feedback, these characteristics of the experimental task are conducive to differential allocation of attention among evidence and therefore are conducive to subjects demonstrating inverse base rate effects in the later transfer phase. Thus, the lack of inverse base rate effect identified in the experiment should generalize to audit practice.

It is unclear why no inverse base rate effect was observed in the analytical procedures context. One possibility is that, regardless of the extent of their accounting knowledge, all subjects knew that ratio differences and financial statement errors were related meaningfully in the analytical procedures context but that symptoms and imaginary diseases were not related meaningfully in the abstract medical diagnosis context. Thus, all subjects may have tried to attend to all items of evidence in the analytical procedures context, and therefore avoided any differential allocation of attention. This explanation would suggest that the inverse base rate effect may not be apparent in meaningful diagnosis contexts in general, which in turn suggests that the decision aids recommended by previous studies (e.g., Medin and Edelson 1988) to counteract the inverse base rate effect may be unnecessary.

References

Abdolmohammadi, M., and A. Wright. 1987. An examination of the effects of experience and task complexity on audit judgments. *The Accounting Review* 62 (January): 1–13.

American Institute of Certified Public Accountants. 1988. *Statement of Auditing Standards No. 56: Analytical Procedures*. New York: AICPA.

Arens, A. A., and J. K. Loebbecke. 1991. *Auditing: An Integrated Approach*. 5th ed. Englewood Cliffs, NJ: Prentice-Hall.

Ashton, A. H. 1991. An empirical analysis of expertise, experience, and error frequency knowledge in auditing. *The Accounting Review* 66 (April): 218–39.

Bedard, J. C., and S. F. Biggs. 1991. Pattern recognition, hypotheses generation, and auditor performance in an analytical task. *The Accounting Review* 66 (July): 622–42.

Begg, I., D. Maxwell, J. O. Mitterer, and G. Harris. 1986. Estimates of frequency: Attribute or attribution? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12 (October): 496–508.

Bickel, P. J., and K. A. Doksum. 1977. *Mathematical Statistics*. San Francisco: Holden-Day.

Bonner, S. E. 1990. Experience effects in auditing: The role of task-specific knowledge. *The Accounting Review* 65 (January): 72–92.

——, and B. L. Lewis. 1990. Determinants of auditor expertise. *Journal of Accounting Research* 28 (Supplement): 1–20.

Butt, J. L. 1988. Frequency judgments in an auditing-related task. *Journal of Accounting Research* 26 (Autumn): 315–30.

Demski, J. S. 1970. Optimizing the search for cost deviation sources. *Management Science* 16 (April): B486–B494.

Elstein, A. S., L. S. Schulman, and S. A. Sprafka. 1978. *Medical Problem Solving: An Analysis of Clinical Reasoning*. Cambridge, MA: Harvard University Press.

Estes, W. K., J. A. Campbell, N. Hatsopoulos, and J. B. Hurwitz. 1989. Base-rate effects in category learning: A comparison of parallel network and memory storage-retrieval models. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15 (July): 556–71.

Fisk, A. D., and W. Schneider. 1984. Memory as a function of attention, level of processing, and automation. *Journal of Experimental Psychology: Learning, Memory and Cognition* 10 (April): 181-97.

Frederick, D. M., and R. Libby. 1986. Expertise and auditors' judgments of conjunctive events. *Journal of Accounting Research* 24 (Autumn): 270–90.

Gluck, M. A., and G. H. Bower. 1988a. Evaluating an adaptive network model of human learning. *Journal of Memory and Language* 27 (April): 116–95.

——, and ——. 1988b. From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General* 117 (September): 227–47.

Hasher, L., and R. T. Zacks. 1979. Automatic and effortful processes in memory. *Journal of Experimental Psychology: General* 108 (September): 356–88.

——, and ——. 1984. Automatic processing of fundamental information. *American Psychologist* 39 (December): 1372–88.

Heiman, V. B. 1990. Auditors' assessments of the likelihood of analytical review explanations. *The Accounting Review* 65 (October): 875–90.

Hintzman, D. L. 1976. Repetition and memory. In *The Psychology of Learning and Motivation 10*, edited by G. H. Bower. New York: Academic Press.

——, G. Nozawa, and M. Irmscher. 1982. Frequency as a nonpropositional attribute of memory. *Journal of Verbal Learning and Verbal Behavior* 21 (April): 127–41.

Jensen, D. L., and J. C. McKeown. 1977. Multiple-step investigations of standard cost variances. *Journal of Business, Finance & Accounting* 4 (Summer): 247–56.

Kinney, W. R., Jr. 1987. Attention-directing analytical review using accounting ratios: A case study. *Auditing: A Journal of Practice & Theory* 6 (Spring): 59–73.

Libby, R. 1985. Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research* 23 (Autumn): 648–67.

——. 1993. The role of knowledge and memory in audit judgment. In *Judgment and Decision Making Research in Accounting and Auditing*, edited by R. H. Ashton and A. H. Ashton. Cambridge, MA: Cambridge University Press.

——, and D. M. Frederick. 1990. Experience and the ability to explain audit findings. *Journal of Accounting Research* 28 (Autumn): 348–67.

——, and H.-T. Tan. 1993. Modeling the determinants of audit expertise. Working paper, Cornell University.

Maki, R. H., and R. S. Ostby. 1987. Effects of level processing and rehearsal on frequency judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13 (January): 151–63.

Medin, D. L., and J. G. Bettger. 1991. Sensitivity to changes in base-rate information. *American Journal of Psychology* 104 (Fall): 311–32.

——, and S. M. Edelson. 1988. Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology: General* 117 (March): 68–85.

——, and M. M. Schaffer. 1978. Context theory of classification learning. *Psychological Review* 85 (May): 207–38.

Naveh-Benjamin, M., and J. Jonides. 1986. On the automaticity of frequency coding: Effects of competing task load, encoding strategy, and intention. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12 (July): 378–86.

Nelson, M. W. 1993. The learning and application of frequency knowledge in audit judgment. Working paper, Cornell University.

Nosofsky, R. M. 1986. Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115 (March): 39–57.

——. 1988. Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14 (October): 700–08.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 825–839

Strategic Auditing for Fraud

Sanford Morton
University of British Columbia

SYNOPSIS AND INTRODUCTION: Two features of current audit practice are striking. First, auditors often use sampling methods to audit *probabilistically*. Second, the probability of auditing is usually *contingent* on information about that item. This essay studies an audit-reporting game in which a manager reports on the privately observed value of the firm. The report may be understated if some of this value is appropriated to the manager's own use. Only a costly audit can directly verify the report and impose penalties on the manager when fraud is discovered. The optimal audit policy for the owner of the firm is completely characterized and is shown to be both probabilistic and contingent: a report is never audited if it is sufficiently large, and is audited with a probability that is strictly less than one, otherwise.

The optimal audit policy is contingent on the manager's report because the report depends on, and so provides some signal of, the firm's value. In contrast to optimal strategies in variance investigation models, the audit policy is also probabilistic. Variance investigation models are based on moral hazard problems, and efficient monitoring is deterministic because the agent has imperfect control over the outcome. In the adverse selection model studied here, the agent has perfect control over the report and would never choose a fraudulent report that will certainly be discovered. This makes probabilistic monitoring efficient.

Because the model yields an exact solution, it can generate some useful comparative statics. As the audit cost decreases, the audit region expands and the amount of misreporting declines. The model also shows how the auditor's information affects the audit strategy and suggests how additional information can be efficiently incorporated into the audit policy.

Key Words: *Auditing, Reporting, Fraud, Adverse selection.*

THE next section motivates the model and describes its place in the audit literature. The second section gives a formal statement of the model. The third section shows that the optimal audit policy is a stratified random audit. Comparative statics are studied in the fourth section. The concluding section discusses implications and generalizations of the model. All proofs are in the Appendix.

With thanks to Kalyan Chatterjee, Barry Cushing, Gerald Feltham, Frank Selto, Joel Sobel, and the editors and referees. They are not responsible for any errors or irregularities.

Submitted September 1989.
Accepted February 1993.

I. Background

The type of audit-reporting game studied here has provided the foundation for several prior studies, beginning with Townsend's (1979) abstract model and continuing with Evans' (1980) analysis of a financial audit and Reinganum and Wilde's (1985) model of a tax audit. However, they did not find the optimal audit policy. Both Townsend and Evans assumed that auditing would be deterministic, leading Fellingham (1980) to speculate that probabilistic audit rules would be an improvement. Townsend proved this was so. (Subsequently, in a three-person model, Baiman et al. [1987] also assume the auditor will use pure strategies.) Similarly, Reinganum and Wilde analyze probabilistic but noncontingent, and then contingent but deterministic audit strategies, and conjecture that a combination of both might be better. This conjecture is verified and a solution is presented here.

Monitoring policies have also been the focus of the variance investigation literature. Baiman and Demski (1980) (see, more recently, Dye [1986] and his references) permit the monitoring policy to be both contingent and probabilistic and find that a deterministic policy is optimal. However, these are moral hazard models in which the auditor is attempting to verify the agent's effort. Deterministic policies appear to be optimal because the agent has only imperfect control over the outcome upon which the investigation is based.[1] In contrast, in an adverse selection or private information reporting model, the agent has complete control over the report and would never make a false report that will be audited with probability one. Therefore, deterministic auditing offers excessive deterrence. The incentive problems presented by reporting seem better modeled as private information, rather than effort-related shirking, since there is no demand for reporting in a pure moral hazard model. Because audits attest to a report, and managers have control over what they say, audits ought generally to be probabilistic.

The gains to randomization in the audit strategy have been noted before. Border and Sobel (1987) conduct a welfare analysis of tax auditing and provide necessary and sufficient conditions for efficient auditing. Because they do not specify an objective function for the principal, their conditions are independent of audit cost and prior information and do not show how these will influence the optimal audit strategy. Mookherjee and Png (1989) do specify an objective function and show that the audit strategy should always be probabilistic, in addition to establishing some monotonicity properties of the reporting strategy. But they do not fully characterize the solution and do not study the effects of cost and prior information. Sanchez and Sobel (1989) analyze a tax audit problem equivalent to the one here and obtain similar results. They are particularly interested in a budget-constrained auditor in order to study a hierarchical problem.

II. Statement of the Model

There are two players in this audit-reporting game, an owner of a firm that generates a random profit and a manager of the firm, who must report on and transfer the profit to the owner. Only the manager directly observes the profit and the resulting value of the firm. To retain the residual, the manager can underreport and transfer less than the full value. The owner can verify the manager's report only with a costly audit, but if an audit uncovers a misreport, the manager can be penalized.

---

[1] An interesting exception is Kanodia (1985). By making the state observable to the agent, he effectively transforms a moral hazard model into adverse selection. He also shows that investigation should be random.

Let x be the random value of the firm, with $F(x)$ its cumulative distribution function on bounded support, $0 \le x \le X$, representing the owners's priors on the value of the firm. It is assumed that the probability density $f(x) \equiv F'(x)$ is positive and continuous. After observing the value of the firm the manager must report $r \in (0,X)$. After observing the report, the owner can conduct an audit at a cost $c > 0$, which will publicly reveal the payoff and any misreport or fraud. If no audit is conducted, the manager retains the residual, $x-r$, and the owner benefits only by the amount reported, r. If an audit is conducted and a fraud is discovered, the manager must return the residual, so the owner benefits by the total payoff, x. The manager is then subject to some additional penalties in proportion to the amount of the fraud, at the penalty rate, L, which accrue to the owner.[2]

For simplicity, the model is static, so that misstatements in prior periods cannot be penalized now, and those not discovered now cannot be penalized in the future. The game has also abstracted from explicit contracting issues between the owner and manager. However, the audit policy itself is an implicit contract, since the owner may choose to audit in a way that permits the manager to supplement his or her base compensation by consuming some part of the value of the firm.[3] In particular, the possibility of selling or renting the firm to the manager at its expected value is ruled out by assuming that the manager has insufficient initial wealth and no access to capital markets, because this would merely shift the audit problem to a lender.

Finally, it is also assumed that both players are risk neutral. (The concluding section discusses the generalization to risk aversion.) The payoffs in the game can be summarized as follows: if the value of the firm is x, the manager chooses to report r and an audit occurs with probability p, then the expected gain to the manager above any base compensation is:[4]

$$u(p,r,x) = (1-p)(x-r) - pL(x-r), \tag{1}$$

and for the owner is:

$$\pi(p,r,x) = (1-p)r + p[x + L(x-r) - c]. \tag{2}$$

It is assumed that the owner is a dominant player and can commit to an audit policy before observing the manager's report. The owner chooses an audit policy, $p(r)$, which specifies the probability that each report will be audited, and the manager responds to this policy by choosing how much to report and transfer, given the risks imposed by this audit policy.[5] The report issued by the manager will generally depend on x, the value of the firm, so the owner must predict the manager's reaction to the audit policy for every possible value of the firm to evaluate the effectiveness of the policy.

---

[2] The punishment policy cannot be a purely contractual matter, but must accord with institutional or social constraints that are external to the employment contract, e.g., the criminal and civil justice system. The constraints assumed here are common in the literature on the economics of crime (Becker 1968; Stigler 1970) and tax auditing (Reinganum and Wilde 1986; Sanchez and Sobel 1989), namely that the "punishment must fit the crime" or that the loss that can be imposed on the manager is some increasing function of the size of the fraud. Linear penalties are the most plausible approximation to the outcome of a complex and uncertain judicial process.

[3] In examining contracting issues in audit models, Border and Sobel (1987) and Mookherjee and Png (1989) have shown that a crucial role for contracting may be to *reward* the manager when an audit reveals a truthful report. This practice is not typically observed and is ruled out here because a manager who has reported correctly would insist on being audited to obtain the reward, and the owner would then lose control of the audit decision.

[4] In the case that $r > x$ and the manager transfers an amount *greater* than the value of the firm, u says the manager is *rewarded* in proportion to L. It will be shown that the manager will never choose this in equilibrium, so it is not necessary to specify a more plausible scenario for this case.

If r(x) is the report the manager issues when observing value x, then this reporting strategy must satisfy,

$$r(x) \in \arg \max_{0 \leq r \leq X} u[p(r),r,x]. \tag{3}$$

Note that by reporting the full value of the firm, $r(x)=x$, the manager can guarantee at least zero expected gain.

Using this constraint, the owner can predict the manager's reaction to the audit policy $p(r)$. An optimal audit policy maximizes the owner's expected gain when the manager's reaction to the audit policy is taken into account.

Definition 1: An optimal audit policy is an audit probability schedule, $p(r)$, which solves the audit problem:

$$\max_{0 \leq p(r) \leq 1} E_x \pi\{p[r(x)],r(x),x\}$$

subject to constraint (3) where E_x is the expectation with respect to x.

Constraint (3) ensures that the reporting strategy the owner forecasts the manager will use will be an optimal reporting strategy for the manager, given the audit policy. For some policies, there may be multiple solutions to constraint (3) because the manager is indifferent among several reports, perhaps including truthful reporting. In such cases, the convention used here is that the manager will be willing to report in whatever manner the owner prefers, as long as the reporting strategy satisfies constraint (3).

III. The Optimal Audit Policy

The owner can choose any audit policy that specifies a probability with which every report will be audited. This includes increasing policies, such as probability proportional to size policies in which the audit probability is proportional to the report, constant or noncontingent policies in which the probability is independent of the report, decreasing policies, and even stratified or discontinuous policies. Among all such policies, the best for the owner in this game is a stratified random audit: a lower-tailed audit policy with a constant probability of audit within the lower audit region. This probability will be strictly less than one, but will be just high enough to deter misreporting within the audit region. However, it will not be efficient for the auditor to completely deter misreporting everywhere, so some reports will never be audited.

---

[5] An alternative extensive game form, in which the owner chooses p after observing r, would employ a perfect Bayes Nash or "sequential" equilibrium and lead to somewhat different results; see Chatterjee and Morton (1990) for such an analysis. However, their formulation essentially removes the owner's commitment ability. Because the owner could choose their equilibrium audit policy but prefers not to, the owner does strictly better in the dominant player game analyzed here.

Proposition 1: An optimal audit policy is given by,

$$p(r)=\begin{cases}\dfrac{1}{1+L} & \text{if } r<r^* \\ 0 & \text{if } r\geq r^*\end{cases} \tag{4}$$

for some $r^* \in (0,X)$.

The optimal audit policy (4) differs from strategies analyzed in prior audit and variance investigation models in being both probabilistic and contingent (exceptions are Kanodia 1985; Sanchez and Sobel 1989). It may be described as a stratified random audit (with two strata) since every report in the lower-tailed audit region, $r<r^*$, will be audited with the same probability, $1/(1+L)$, and no report in the higher stratum, $r\geq r^*$, will ever be audited. The audit policy is lower-tailed because the manager benefits by underreporting the value of the firm, so low reports are more likely to be understated in relation to the actual value.[6] The manager's expected payoff (1) is $[1-p(1+L)](x-r)$, so reports below the stratification point, r^*, are audited with the probability $1/(1+L) <1$ that is just large enough to yield an expected gain to the manager of zero.

The manager's reporting behavior under this audit policy depends on the observed value of the firm. If x is large enough, $x>r^*$, the manager will report only r^* and retain the rest. According to the audit policy, (4), this report will not be audited.[7] The manager will not report less than this since a lower report will be audited just often enough to reduce the expected gain to zero. However, if $x\leq r^*$, the manager cannot avoid a random audit (except by transferring to the owner an amount greater than the actual value of the firm). The manager will be indifferent among all reports less than x and will be willing to report the full value of the firm, $r(x)=x$. Therefore, an optimal reporting strategy under the audit policy (4) is:

$$r(x)=\begin{cases}x & \text{if } x<r^* \\ r^* & \text{if } x\geq r^*.\end{cases} \tag{5}$$

The owner audits reports in the audit region, $r<r^*$, with sufficient probability to deter misstatement from those reports. Misreporting occurs only outside the audit region, in particular, at r^*, and is never audited. Therefore, the optimal audit policy never discovers a misreport. Nevertheless, if the audit region is not empty, the optimal audit policy does partially deter fraud.

Since fraud is never discovered, it follows that even if the manager's penalties did not accrue to the owner, as is plausible for imprisonment and other nonmonetary penalties, the optimal audit policy would be the same. In this case, the manager's payoffs and the constraints in the audit problem would be unchanged, but the owner's payoffs (2) cannot increase since a source of revenue has been removed. Therefore, the owner could never achieve a strictly higher value in the audit problem. But under policy (4), the owner receives the same value since fraud is never discovered and the manager is never penalized, so policy (4) is optimal in both cases.

[6] This results from modeling underreporting as the basic incentive problem. If the manager's incentives were to overreport (to gain greater bonus compensation or to enhance future career prospects) then the optimal audit policy would be upper-tailed. Chatterjee and Morton (1990) analyze an overreporting problem.

[7] In fact, r^* must not be audited in a continuous model. Otherwise, when $x>r^*$, the manager will attempt to report as little above, but strictly greater than, r^* as possible, and a solution to the manager's reporting problem (3) would not exist.

IV. Comparative Statics

This section examines how the optimal audit policy depends on the parameters of the model, including the audit cost, the penalty rate, and the owner's probability beliefs about the value of the firm.

Given the form of $p(r)$ specified by policy (4), and the corresponding reporting strategy in policy (5), the owner's expected profits can be written as:

$$\prod(r^*,c)\equiv\int_0^{r^*}\left(x-\frac{c}{1+L}\right)dF(x)+r^*[1-F(r^*)]. \tag{6}$$

The stratification point, r^*, will be chosen to maximize policy (6). Since $F(x)$ is continuous, Π is continuous in r^* on the compact interval $(0,X)$, so an optimal r^* exists. This establishes existence of an optimal audit policy.

The simplest example of the calculation of the stratification point occurs when x is distributed uniformly over $[0,X]$; in this case,

$$\prod(r^*,c)=\frac{r^*\left(X-\frac{c}{1+L}\right)-\frac{r^{*2}}{2}}{X}.$$

Differentiating with respect to r^* shows that the optimal stratification point will be,

$$r^*=\max\left(X-\frac{c}{1+L},0\right)\cdot$$

The solution depends on a comparison between the owners priors, determined by the upper bound, X, and the probability weighted audit cost, $c/(1+L)$. If auditing becomes more costly (or the penalty rate on the manager decreases), then the audit region will shrink. If X increases and the owner thinks it more likely that the payoff was large, then the audit region expands. If the audit cost is sufficiently high in relation to the owner's priors, $c\geq X(1+L)$, then it is most efficient not to audit at all, so $r^*=0$, but r^* is always less then X, so it is never optimal to audit all reports.

Most of these results can be established more generally. First, consider a change in the audit cost. For an audit policy of the form (4), a change in c can only change r^*. Let $r^*(c)$ be the optimal stratification point when the audit cost is c:

$$r^*(c)\in\underset{r^*}{\arg\max}\prod(r^*,c).$$

One would expect that the amount of auditing increases as the audit cost decreases.

Corollary 1: The optimal stratification point, $r^*(c)$, is nonincreasing in c. The owner's optimal profits, $\Pi[r^*(c),c]$, are nonincreasing in c.

When auditing becomes less costly, more auditing will occur, but the increase in the amount of auditing takes a special form. The *extensiveness* of auditing will increase, in the sense that more reports will become subject to audit. The audit region expands and additional misstatement is deterred. On the other hand, the *intensiveness* of the audit is unchanged, in the sense that the probability that the report will be audited, given that it lies in the audit region, is independent of c and depends only on L.

It also follows that the optimal stratification point is nondecreasing in the penalty rate, L. This is because c enters $\Pi(r^*, c)$ only through the probability weighted audit cost, $c/(1+L)$. Therefore, an increase in L will have the same effect as a decrease in c, namely, an increase in r^*. Unlike the audit cost, however, an increase in the penalty rate will also decrease the intensiveness of auditing, since the audit probability required to deter misreporting, $1/(1+L)$, is lower.

Next, consider the limiting cases of very low or very high audit cost. It is never optimal to audit exhaustively, as long as the audit cost is positive. On the other hand, if the audit cost is sufficiently high, it will be optimal to never audit. These results rely on the earlier assumptions of bounded support for x and strictly positive likelihood of x throughout its support.

Corollary 2: For any positive audit cost, it is never optimal to audit all reports. For large enough audit cost, it is optimal never to audit.

Finally, consider how the owner's probability beliefs affect the audit strategy. If the owner receives new information that makes higher values of the firm more likely (a first order stochastic shift in the distribution of x) then lower reports that were reasonable before may now be less reasonable and more likely to be misstated. In this case, one might expect that the lower-tailed audit region will expand. In addition, because the firm has a higher expected value, the owner's profits ought to increase.

This conjecture about profits is correct, but the determination of the audit region is more complicated. Differentiating policy (6) with respect to r^* yields the first order condition:

$$\Pi_{r^*}(r^*, c) = [1-F(r^*)]\left(1-\frac{c}{1+L}\,\frac{f(r^*)}{1-F(r^*)}\right)=0. \tag{7}$$

The distribution determines the stratification point through its *hazard function*, $f/(1-F)$, which is the conditional likelihood that x is equal to r^* given that it is no less than r^*. If the distribution shifts, then the change in the stratification point depends on the corresponding shift in the hazard function. First order stochastic dominance is not sufficient to predict the shift in the hazard function, since it does not determine the change in the density, f. It must also be that the likelihood of firm values around the stratification point does not increase too much.

Corollary 3: (1) Suppose the owner's distribution changes from F to G, and G first order stochastically dominates F [i.e., for all x, $G(x) \leq F(x)$]. Then the owner's expected optimal profits will (weakly) increase. (2) Suppose further that both hazard functions, $f/(1-F)$ and $g/(1-G)$, are strictly increasing, where $g(x)=G'(x)$, and also that $g(r_F^*) \leq f(r_F^*)$, where r_F^* is the optimal stratification point under F. Then $r_F^* \leq r_G^*$, where r_G^* is the optimal stratification point given G.

The proof shows that increasing hazard functions guarantee that the first order condition (7) will characterize an interior solution. The proof also shows that even if higher firm values become more likely, but values near the stratification point also become more likely, the audit region may contract.

V. Conclusion

This essay has explored a simple audit game, to study the interaction between the audit and reporting strategies. Although this interaction was potentially complex, the optimal audit policy was found to take a simple form, a stratified random audit. This section discusses some implications and generalizations of the model.

Under the optimal audit policy, misreporting is never discovered. In this model, whenever it is worthwhile to audit some report at all, it is optimal to audit with a high enough probability to deter fraud completely from that report. In part, this is because of the assumption that only irregularities or intentional fraud, rather than simple errors or unintentional mistakes, will cause misstatement. If purely random and nonstrategic mistakes were incorporated into the model (and were frequent enough to merit the audit cost to recover them), then policy (4) will likely require revision. However, if these errors are indeed purely random and uncontrollable by the manager, then there would be no strategic reaction to consider and the decision to audit would be a simple matter of comparing expected benefits and costs. Thus, the revision may take a simple form: audit a report with probability one if the expected error is greater than the audit cost; otherwise, follow policy (4). If the errors were expected to be small, then policy (4) would not require revision, yet misreports would occasionally be discovered.[8]

The assumption of risk neutrality was important in this model because the analytical methods used here cannot easily be extended to nonlinear preferences. If the manager were risk averse, then, at a minimum, the amount of auditing specified by policy (4) could be decreased—since policy (4) deters a risk-neutral manager, it provides excessive deterrence to a risk-averse manager. However, there may be more fundamental changes. Matthews (1983) was able to extend Myerson's (1981) auction analysis (whose analytical techniques were similar to those employed here) to the case of constant absolute risk-averse bidders and showed that risk should be imposed on the bidders to extract greater surplus. In this model, risk is not imposed on the manager because fraud is never discovered; to impose risk on the manager requires that fraud occasionally be discovered, so that policy (4) should be decreased even further. The optimal policy in this case will depend on the risk attitude of the manager, which is generally not known to the auditor. Although policy (4) would lead to excessive auditing of a risk-averse manager, it may nevertheless be a sensibly conservative approach under uncertainty about the manager's risk aversion.

Some conjectures can also be made about the effects of perturbing the objective of the audit, which, in this model, was to minimize misreporting net of audit cost. These objectives seem to be natural in a setting where a manager is to be deterred from fraud, but if the audit is conducted by an independent auditor who is distinct from the owner, the auditor may have additional concerns. Although the immediate task is the expression of an opinion, an auditor's ultimate interests are legal liability and reputation. It is easy to imagine that deterring misreporting can play some role in these goals, but there may be other purposes that will have some effect on the audit policy.[9] Some idea of how these other purposes may affect the audit policy can be gained by incorporating a "budget" constraint into the audit problem, to the effect that the expected amount of auditing must be no less than some budget level:

$$E_x \hat{p}[r(x)] \geq B.$$

---

[8] For an approach to the audit-reporting game in which it is not optimal to deter misreporting completely, see Chatterjee and Morton (1990).

[9] There also will be important contractual issues when the auditor is distinct from the owner, as in Baiman et al. (1987).

This constraint can proxy for other purposes that may require some additional auditing than is otherwise provided by policy (4), such as industry or legal standards that require some minimum amount of auditing.

Linear constraints of this sort have been examined by Samuelson (1984) and Sanchez and Sobel (1989), who have shown that the solution will still be a step function, but may now require three steps:

$$p(r)=\begin{cases}\dfrac{1}{1+L} & \text{if } 0\leq r<r_1^* \\ k & \text{if } r_1^*\leq r<r_2^* \\ 0 & \text{if } r_2^*\leq r\end{cases}$$

for some $k\in[0,1/(1+L)]$.[10] An intermediate step, k, may be required if policy (4) does not already achieve the minimum amount of auditing. In this case, the manager may sometimes choose to misreport $r_1^*<x$, since k is not large enough to deter fraud and the manager will prefer the lowest of the reports that offer the same risk, k, of discovery. Although a misreport will occasionally be discovered by an audit, Proposition 1 shows this is not optimal purely from the owner's considerations.

A final comment concerns the use of additional information in the optimal audit policy. There will often be much information in the audit environment on which to base the audit decision, such as that obtained from prior review. The model suggests when this information will be relevant and how it should be incorporated into the audit policy: it will be relevant when it affects the owner's priors on the value of the firm, and it will be used to expand or contract the audit region. A common prior review practice is analytical review, which seeks to gauge the reasonableness of the report in light of prior information. The model suggests that such information be channeled into forming priors over the value of the firm's payoff and then be used to choose the extensiveness of auditing, while holding intensiveness constant.

Appendix

Proof of Proposition 1

To simplify the audit problem, it is useful to first establish two qualitative features of optimal audit and reporting strategies. First, the owner can restrict attention to audit probabilities that never exceed $1/(1+L)$, the lowest probability that can deter misreporting. Second, because of this, the manager will never overreport. Other audit policies, which differ on sets of measure zero or for off-equilibrium reports, may do as well but cannot do better.

Lemma 1: If $p(r)$ is an optimal audit strategy and $r(x)$ is the corresponding reporting behavior from constraint (3), then for almost every x, $p[r(x)]\leq 1/(1+L)$ and $r(x)\leq x$.

Proof: Because $u[p(r),r,x]=[1-p(r)(1+L)](x-r)$, by choosing $r=x$ the manager can always guarantee $u=0$. The manager will strictly prefer to overreport, $r>x$, if and only if $p(r)>1/(1+L)$ for some $r>x$. But using payoffs (1) and (2), the owner's payoffs can be written $\pi(p,r,x)=x-pc-u(p,r,x)$, so the owner wishes to minimize both the probability of audit and the manager's expected payoff. By lowering all audit probabilities greater than $1/(1+L)$, the owner can reduce both expected audit cost and the manager's payoff for those values of x for which $r(x)>x$, without altering the behavior of values of x for which $r(x)\leq x$. Q.E.D.

---

[10] This result can also be seen by modifying the proof of Lemma 4 in the Appendix. The constraint requires that when an interior interval is lowered, then some other interval must be raised to preserve B. If there is at least one other interior interval, this can always be done. Each iteration still reduces the number of intervals by one. When there is only one interior interval left, the process may have to terminate, since any further improvement may violate the constraint.

The remainder of the proof uses revelation techniques developed by Myerson (1981) for an auction model. The audit policy, $p(r)$, is an example of an *allocation* mechanism or game chosen by the owner that the manager plays by choosing an optimal reporting strategy, $r(x)$. Every such allocation mechanism can always be represented by a *revelation* mechanism, which is a two-stage game with both private and public messages: first, the manager confidentially reveals a value, χ, of the private information to the owner; then the revelation mechanism maps this into a payoff or allocation consisting of a probability of audit $\hat{p}(\chi)$, and a publicly issued report, $\hat{r}(\chi)$. Applying Myerson's "revelation principle" to this environment results in,

Lemma 2: For any audit policy, a revelation mechanism can always be found that will induce the manager to truthfully reveal the private information, $\chi=x$. Further, the payoffs will always be the same as those resulting from the original audit policy.

Proof: Consider any audit policy, $p(r)$, and corresponding optimal reporting strategy, $r(x)$ defined by constraint (3). Define a revelation mechanism from them by,

$$\hat{r}(x)=r(x),\hat{p}(x)=p[r(x)]. \tag{8}$$

The manager can get the same payoff in this revelation game by revealing truthfully, $\chi=x$, as in the original audit game. Further, the manager will not prefer to reveal any other value in the revelation game: if the manager prefers to reveal $\chi\neq x$, it is because the $\hat{p}(\chi),\hat{r}(\chi)$ payoff is preferred to $\hat{p}(x),\hat{r}(x)$. If this is so, the manager would have chosen $r(\chi)$ in the audit game rather than $r(x)$, which contradicts constraint (3). Q.E.D.

In applying the revelation principle in a reporting model, it is important to note that the publicly issued report is not itself a confidential revelation. If it were, then truthful revelation would require that $r(x)=x$, or that the manager never commit fraud. Since some audit policies do permit fraud, it is necessary to allow misreporting. Revelation mechanisms permit this by distinguishing the public report, $\hat{r}$, which determines the amount of fraud, from the confidential revelation, χ, made to the owner, so that truthful revelation can allow for public misreporting and fraud. Note also that even if the owner is sure that the confidential revelation is true and the public report is false, $\chi=x\neq r(x)$, penalties cannot be imposed unless an audit verifies that this is so.

There is no claim that revelation mechanisms are ever used in practice, only that they can mimic any other mechanism that might be used. Revelation functions are easier to analyze in this model because they replace the compound function $p[r(x)]$ with the simpler pair of functions $\hat{p}(x),\hat{r}(x)$. Therefore, the proof proceeds by first finding the truthful revelation mechanism that is best for the owner. With this in hand, the final step in the proof is to use constraint (8) to translate back to a $p(r)$ audit policy that is equivalent to that mechanism.

Definition 2: An optimal truthful revelation mechansim will solve:

$$\max_{\hat{p}(x),\hat{r}(x)} \quad E_x\pi[\hat{p}(x),\hat{r}(x),x]$$

subject to:

$$x\in \underset{\chi}{\arg\max}\, u[\hat{p}(\chi),\hat{r}(\chi),x] \tag{9}$$

$$u[\hat{p}(x),\hat{r}(x),x]\geq 0,$$

$$0\leq\hat{p}(x)\leq\frac{1}{1+L}, \text{ and } 0\leq\hat{r}(x)\leq x.$$

The first constraint (9) ensures that the mechanism is truthful, the manager will prefer to reveal $\chi=x$. The second is a participation constraint that guarantees that the manager will always obtain a nonnegative expected payoff. It was not necessary in the earlier audit problem because the manager could always report $r(x)=x$ to assure this. The last constraints are just Lemma 1, above.

The next step in the proof is to express constraint (9) in a convenient "envelope" form condition (10).

Lemma 3: If $\hat{p}(x), \hat{r}(x)$ form an optimal revelation mechanism, they satisfy constraint (9) if and only if both of the following hold.

$$u[\hat{p}(x),\hat{r}(x),x]=u[\hat{p}(x'),\hat{r}(x'),x']+\int_{x'}^{x} 1-\hat{p}(t)(L+1)dt \tag{10}$$

$$\hat{p}(x) \quad \text{is nonincreasing.} \tag{11}$$

Proof: If constraint (9) holds, then $\hat{p}(x),\hat{r}(x)$ is preferred when the manager observes x to $\hat{p}(x'),\hat{r}(x')$, the preferred choice when x' is observed. This is equally true when x and x' are interchanged. Combining these two preference relations,

$$u[\hat{p}(x'),\hat{r}(x'),x]-u[\hat{p}(x'),\hat{r}(x'),x'] \le u[\hat{p}(x),\hat{r}(x),x]-u[\hat{p}(x'),\hat{r}(x'),x']$$
$$\le u[\hat{p}(x),\hat{r}(x),x]-u[\hat{p}(x),\hat{r}(x),x'].$$

After substituting payoff (1) and rearranging, this becomes,

$$[1-\hat{p}(x')(1+L)](x-x') \le u[\hat{p}(x),\hat{r}(x),x]-u[\hat{p}(x'),\hat{r}(x'),x']$$
$$\le [1-\hat{p}(x)(1+L)](x-x').$$

If $x>x'$, this implies $\hat{p}(x') \ge \hat{p}(x)$, so condition (11) holds. The last inequality also establishes that the indirect utility function $u[\hat{p}(x),\hat{r}(x),x]$ is Lipschitzian in x and therefore equal to the integral of its derivative (Kolmogorov and Fomin 1970, 340–42). Dividing by $x-x'$ shows that the derivative is equal to $1-\hat{p}(x)(L+1)$ whenever it exists. Therefore condition (10) holds.

Conversely, since $u(p,r,x)$ is linear in its third argument,

$$u[\hat{p}(x'),\hat{r}(x'),x]=u[\hat{p}(x'),\hat{r}(x'),x']+\int_{x'}^{x} 1-\hat{p}(x')(L+1)dt.$$

Subtract this from condition (10) to obtain,

$$u[\hat{p}(x),\hat{r}(x),x]-u[\hat{p}(x'),\hat{r}(x'),x]=\int_{x'}^{x}[\hat{p}(x')-\hat{p}(t)](L+1)dt.$$

This is nonnegative if $\hat{p}(x)$ is nonincreasing. Thus, constraint (9) holds. Q.E.D.

The envelope condition (10) can now be incorporated into the owner's objective function by substitution. First, $E_x u[\hat{p}(x),\hat{r}(x),x]$ can be written as:

$$\int_0^X u[\hat{p}(x),\hat{r}(x),x]dF(x)=u[\hat{p}(x),\hat{r}(x),x]F(x)\Big|_0^X-\int_0^X F(x)du[\hat{p}(x),\hat{r}(x),x]$$
$$=\int_0^X [1-F(x)][1-\hat{p}(x)(L+1)]dx,$$

where the first step follows from integration by parts, since both u and F are absolutely continuous. The second step follows from condition (10) and $u[\hat{p}(0),\hat{r}(0),0)]\equiv 0$. Then, because $\pi=x-cp-u$, the owner's objective function can be written as:

$$\int_0^X \pi[\hat{p}(x),\hat{r}(x),x]dF(x)=\int_0^X [x-c\hat{p}(x)]f(x)-[1-F(x)][1-\hat{p}(x)(L+1)]dx$$
$$=\int_0^X \hat{p}(x)(L+1)\left\{[1-F(x)]-\frac{c}{1+L}f(x)\right\}dx. \tag{12}$$

This establishes the following.

Lemma 4': An optimal truthful revelation mechanism will consist of $\hat{p}(x)\in[0,1/(1+L)]$ and $\hat{r}(x)\in(0,x)$, which maximizes expression (12) subject to conditions (10) and (11).

The advantage of substituting condition (10) into the owner's objective function in this way is that $\hat{r}(x)$ has been removed from expression (12), which now depends only on $\hat{p}(x)$. One can then choose $\hat{p}(x)$ to maximize expression (12) subject to condition (11) and then choose $\hat{r}(x)$ to satisfy condition (10) to obtain a truthful revelation mechanism.

If the coefficient on $\hat{p}(x)$ in the integrand of expression (12) is strictly decreasing, the optimal $\hat{p}(x)$ will be the pointwise maximum of expression (12): as large as possible when the coefficient is positive, as small as possible otherwise. The resulting $\hat{p}(x)$ will then satisfy condition (11). (Myerson [1981] called this the "regular" case.) However, if this coefficient is increasing, this pointwise maximum may violate condition (11) and some other procedure must be used. The following lemma deals with this possibility. In this lemma, $m(x)$ is the coefficient,

$$(L+1)\left\{[1-F(x)]-\frac{c}{1+L}f(x)\right\}.$$

Lemma 5: If $m(x)$ is piecewise (absolutely) continuous, then the problem,

$$\max_{\hat{p}(x)} \int_0^X \hat{p}(x)m(x)dx$$

subject to:

$$0 \leq \hat{p}(x) \leq \frac{1}{1+L},$$

and condition (11) is solved by:

$$\hat{p}(x)=\begin{cases}\dfrac{1}{1+L} & \text{if } x=x^* \\ 0 & \text{if } x\geq x^*\end{cases} \tag{13}$$

for some $x^* \in (0,X)$.

Proof: Let $q(x)$ be any feasible nonincreasing function. By the Lebesgue bounded convergence theorem (Kolmogorov and Fomin 1970, 303), any such function can be arbitrarily closely approximated by a nonincreasing step function that takes on finitely many distinct values, $q_i, i=1,\ldots,n$, so assume $q(x)$ is such a step function. It will be shown that $q(x)$ can be dominated in the problem by a function of the form of condition (13). Then the convergence theorem implies that this is true of all nonincreasing functions as well.

Let $A_i=[x|q(x)=q_i]$. Since $q(x)$ is nonincreasing, the set of all A_i forms a partition of $[0,X]$ consisting of connected intervals. Then let,

$$M_i=\int_{A_i} m(x)dx,$$

so that,

$$\int_0^X q(x)m(x)dx=\sum_i q_i M_i.$$

Take any interior interval, A_{i^*}, not containing 0 or X, and construct an improved, feasible function by raising q_{i^*} to q_{i^*-1} if $M_i^*>0$, or lowering it to q_{i^*+1} if $M_{i^*}\leq 0$. The new function is still nonincreasing and is an improvement because its value on A_{i^*} has been increased (decreased) when the corresponding coefficient is positive (negative). The number of intervals has been reduced by one.

After finitely many repetitions of this process, the resulting function will have only two intervals, A_1 and A_2, where the first is the lower interval and the second is the higher interval. If $M_1>0$, increase the function in that interval to $1/(1+L)$; otherwise, decrease it to q_2. If the function was increased, then perform an analogous construction on A_2: if $M_2\leq 0$, lower the function in that interval to 0; otherwise, raise it to $1/(1+L)$. If $M_1\leq 0$, there is now only one interval. So if,

$$\int_0^X m(x)dx>0,$$

increase the function to $1/(1+L)$ on the combined interval; otherwise, lower it to 0. In each of these constructions, the objective function has been improved, since the function has been increased when the coefficient is positive and decreased when it is negative. This shows that some $\hat{p}(x)$ of the form condition (13) dominates any feasible $q(x)$. Q.E.D.

To conclude the proof of the proposition, the $\hat{r}(x)$ required to make the revelation mechanism truthful is found by substituting condition (13) into condition (10), resulting in:

$$\begin{cases} 0 & \text{if } x<x^* \\ x-\hat{r}(x) & \text{if } x\geq x^* \end{cases} = \max(x-x^*,0).$$

Thus for $x\geq x^*$, $\hat{r}(x)=x^*$; for $x<x^*$, $\hat{r}(x)$ can be chosen to be anything, in particular, $\hat{r}(x)=x$. As was discussed in the text, an audit policy that generates these revelation functions is,

$$p(r)=\begin{cases} \dfrac{1}{1+L} & \text{if } r<r^* \\ 0 & \text{if } r\geq r^*, \end{cases}$$

where $r^*=x^*$. This proves Proposition 1.

Proof of Corollary 1

For the first part of the corollary, because $r^*(c)$ is optimal when the audit cost is c, $\Pi(r^*,c)$ cannot be greater for any other stratification point, such as $r^*=r^*(c')$. Therefore, for every c and c', $\Pi[r^*(c),c]\geq\Pi[r^*(c'),c]$. The converse holds as well when c and c' are interchanged: $\Pi[r^*(c'),c']\geq\Pi[r^*(c),c']$. Combining these two inequalities,

$$\begin{aligned}\prod[r^*(c),c]-\prod[r^*(c),c'] &\geq \prod[r^*(c),c]-\prod[r^*(c'),c'] \\ &\geq \prod[r^*(c'),c]-\prod[r^*(c'),c'].\end{aligned}$$

After substitution, the extreme inequality becomes,

$$\frac{c'-c}{1+L}\int_0^{r^*(c)} dF(x)\geq\frac{c'-c}{1+L}\int_0^{r^*(c')} dF(x).$$

If $c'>c$, then $F[r^*(c)]\geq F[r^*(c')]$. Because F is nondecreasing it follows that $r^*(c')\leq r^*(c)$.

To prove the second part of the corollary, the same reasoning shows,

$$\begin{aligned}\prod[r^*(c),c]-\prod[r^*(c'),c'] &\geq \prod[r^*(c'),c]-\prod[r^*(c'),c'] \\ &= \frac{c'-c}{1+L}F[r^*(c')],\end{aligned}$$

so if $c'>c$, then $\Pi(r^*(c),c]\geq\Pi(r^*(c'),c')$. Q.E.D.

Proof of Corollary 2

For the first part of the corollary, it must be shown that for every $c>0$, there is an $r^*<X$ such that $\Pi(r^*,c)>\Pi(X,c)$. After substitution and rearrangement, this is equivalent to,

$$\frac{c}{1+L}>\int_{r^*}^{X}\frac{dF(x)}{1-F(x)}-r^*.$$

The integral on the right is the truncated mean, $E(x|x\geq r^*)$, which lies between r^* and X. Since the left-hand side is positive, and X is finite, r^* can always be chosen close enough to X to satisfy this inequality.

For the second part, it must be shown that if c is large enough, then for all $r^*>0, \Pi(r^*,c)<\Pi(0,c)=0$. This is equivalent to,

$$\int_0^{r^*}\frac{dF(x)}{F(r^*)}-r^*+\frac{r^*}{F(r^*)}<\frac{c}{1+L},$$

so it must be shown that the left-hand side is bounded for $r^*\in(0,X)$. The integral on the left is the truncated mean, $E(x|x<r^*)<r^*$, so only the third term might be unbounded, in particular at $r^*=0$. Using L'Hôpital's rule,

$$\lim_{r^*\to 0} \frac{r^*}{F(r^*)}=\frac{1}{f(0)}.$$

Consequently, since $f(0)>0$, the left-hand side will be bounded, and c can be chosen large enough to satisfy the inequality. Q.E.D.

Proof of Corollary 3

To prove the first part, define the owner's profits under the distribution G as,

$$\prod(r^*|G)=\int_0^{r^*}\left(x-\frac{c}{1+L}\right)dG(x)+r^*[1-G(r^*)],$$

and let r_G^* be an optimal stratification point given G:

$$r_G^* \in \operatorname*{arg\,max}_{r^*} \prod(r^*|G).$$

Similarly define $\Pi(r^*|F)$ and r_F^*. Because $\Pi(r_F^*|F)\geq\Pi(r_G^*|F)$,

$$\begin{aligned}\prod(r_G^*|G)-\prod(r_F^*|F)&\geq\prod(r_G^*|G)-\prod(r_G^*|F)\\ &=\int_0^{r_G^*}x\,d[G(x)-F(x)]-r_G^*[G(r_G^*)-F(r_G^*)]-\frac{c}{1+L}[G(r_G^*)-F(r_G^*)]\\ &=-\int_0^{r_G^*}G(x)-F(x)dx-\frac{c}{1+L}[G(r_G^*)-F(r_G^*)],\end{aligned}$$

where the last equality follows from integration by parts. If $F(x)\geq G(x)$, for all x, then the last equality is nonnegative, and so $\Pi(r_F^*|F)\leq\Pi(r_G^*|G)$.

To prove the second part, if $r_F^*=0$, the claim is necessarily true, so assume $r_F^*>0$. Because both hazard rates are strictly increasing, there can be at most one solution to the necessary first order condition, and that solution must be the optimal stratification point. The assumptions $G(r_F^*)\leq F(r_F^*)$ and $g(r_F^*)\leq f(r_F^*)$ imply,

$$\frac{g(r_F^*)}{1-G(r_F^*)}\leq\frac{f(r_F^*)}{1-F(r_F^*)}.$$

Since r_F^* is optimal given F, the right-hand side is equal to $(1+L)/c$. Because $g/(1-G)$ is increasing, the value of r^* that makes the left-hand side equal to $(1+L)/c$ must be higher. Therefore, $r_G^*\geq r_F^*$. Q.E.D.

References

Baiman, S., and J. Demski. 1980. Economically optimal performance evaluation and control systems. *Journal of Accounting Research* 18 (Supplement): 184–220.

——, J. H. Evans, and J. Noel. 1987. Optimal contracts with a utility-maximizing auditor. *Journal of Accounting Research* 25 (Autumn): 217–44.

Becker, G. 1968. Crime and punishment: An economic approach. *Journal of Political Economy* 76 (March/April): 169–217.

Border, K. C., and J. Sobel. 1987. Samurai accountant: A theory of auditing and plunder. *Review of Economic Studies* 54 (4): 525–40.

Chatterjee, K., and S. Morton. 1990. Strategic audit policies without commitment. Working paper, Pennsylvania State University, University Park, PA.

Dye, R. 1986. Optimal monitoring policies in agencies. *Rand Journal of Economics* 17 (Autumn): 339–50.

Evans, J. H. 1980. Optimal contracts with costly conditional auditing. *Journal of Accounting Research* 18 (Supplement): 108–28.

Fellingham, J. C. 1980. Discussion of optimal contracts with costly conditional auditing. *Journal of Accounting Research* 18 (Supplement): 129–39.

Kanodia, C. 1985. Stochastic monitoring and moral hazard. *Journal of Accounting Research* 23 (Spring): 175–93.

Kolmogorov, A. N., and S. V. Fomin. 1970. *Introductory Real Analysis*. New York: Dover.

Matthews, S. 1983. Selling to risk averse buyers with unobservable tastes. *Journal of Economic Theory* 30: 370–400.

Mookherjee, D., and I. Png. 1989. Optimal auditing, insurance, and redistribution. *Quarterly Journal of Economics* 104 (May): 339–415.

Myerson, R. B. 1981. Optimal auction design. *Mathematics of Operations Research* 6 (February): 58–73.

Reinganum, J., and L. Wilde. 1985. Income tax compliance in a principal-agent framework. *Journal of Public Economics* 26 (February): 1–18.

——, and ——. 1986. Equilibrium verification and reporting policies in a model of tax compliance. *International Economic Review* 27 (October): 739–60.

Samuelson, W. 1984. Bargaining under asymmetric information. *Econometrica* 52 (July): 995–1006.

Sanchez, I., and J. Sobel. 1989. Hierarchical design and enforcement of income tax policies. Working paper, University of California, San Diego, CA.

Stigler, G. 1970. The optimum enforcement of laws. *Journal of Political Economy* 78 (May/June): 526–36.

Townsend, R. 1979. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory* 21 (October): 265–93.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 840–855

The Timing of Asset Sales and Earnings Manipulation

Eli Bartov
New York University

SYNOPSIS AND INTRODUCTION: This study presents an empirical examination of whether managers manipulate earnings through the *timing* of income recognition from disposal of long-lived assets and investments (hereafter assets). Since managers can often choose the period during which an asset will be sold, and since the principle of acquisition cost underlying the accounting valuation of assets implies that changes in the market value of an asset between acquisition and sale are reported in the period of sale, it follows that there are opportunities for managers to manipulate earnings through the timing of asset sales at relatively low cost.

Two common explanations for earnings manipulations are examined: the earnings-smoothing and the debt-equity hypotheses.[1] The earnings-smoothing hypothesis predicts that earnings are manipulated to reduce fluctuations around some level that is considered normal for the firm (see, e.g., Ronen and Sadan 1981, 6). The debt-equity hypothesis suggests a positive relation between a firm's debt-equity ratio and managers' choice of earnings-enhancing activities (see, e.g., Watts and Zimmerman 1986, 200–21).[2]

The findings are consistent with the timing of asset sales by managers so that the recognized income from these sales smooths intemporal earnings changes and mitigates accounting-based restrictions in bond cov-

---

[1] A third explanation for earnings management, the bonus-plan hypothesis, is not explored in detail because of a lack of available data; however, an attempt is made to control for its effects. The bonus-plan hypothesis assumes that managers maximize their compensation through earnings manipulation and suggests that managers reduce earnings when these are outside the lower or upper bound of the plan and enhance earnings otherwise (see, e.g., Healy 1985). The term "earnings manipulation" typically has a negative connotation. In this article, however, earnings manipulation is not used in a pejorative sense.

[2] Underlying this debt-equity hypothesis is the assertion of contracting theory that managers act to relax accounting-based bond covenants by earnings manipulations, as well as the assumption that leverage is a proxy for the existence and proximity of these covenants. This is further discussed later.

I am indebted to Ray Ball, Joshua Livnat, Ross Watts, Paul Zarowin, Jerry Zimmerman, and to participants in workshops at the University at Buffalo and University of Rochester for their valuable comments. I also acknowledge the editor, James McKeown (the associate editor), and two anonymous reviewers for many useful suggestions. I thank the John M. Olin Foundation for financial support. This work has been presented at the 1992 annual meeting of the American Accounting Association.

Submitted October 1990.
Accepted April 1993.

enants. The results also show that the earnings-smoothing and the debt-equity effects are incremental; after controlling for one effect the other still exists. Finally, sensitivity analysis showing that these results are robust to specifications that consider asset sale levels, bonus plans, possible tax effects, and financial soundness suggests that the findings are not attributable to misspecified statistical tests.

The results of this study may be relevant to the ongoing debate about dropping the principle of acquistion cost underlying the accounting valuation of assets in favor of the principle of current cost (see, e.g., *The Wall Street Journal* 1990b). While the former is considered reliable, the latter is said to be vulnerable to earnings manipulations because net income is materially affected by management opinion on the sale values of the assets involved (see, e.g., Kieso and Weygandt 1987, 35). The evidence in this study suggests that, despite its reliability, managers take advantage of the acquisition-cost principle to manipulate earnings. This should be considered when evaluating the advantages and disadvantages of acquisition-cost accounting relative to current-cost accounting.

Key Words: *Earnings manipulations, Earnings smoothing, Debt-equity hypothesis, Asset sales.*

Data Availability: *Data can be retrieved from sources identified in the article.*

THE remainder of the article is organized into six sections. The motivation for this study is explained in section I, and empirical implications of the earnings-smoothing and debt-equity hypotheses are formulated in section II. The research design is outlined in section III, the data are described in section IV, empirical results are presented in section V, and the final section contains a summary of the findings.

I. Motivation

A central question in financial accounting is whether managers manipulate reported earnings. The business community generally concedes that earnings manipulation in financial reporting is pervasive. For example, *Fortune* (1989, 193) quotes the editor of the *Quality of Earnings Report*, a respected newsletter, as saying, "Scrupulous honesty in financial reporting is at an all-time low. To most companies today, managing earnings is almost irresistible." In the academic literature, however, the evidence of earnings manipulation is suggestive but not conclusive (see, e.g., Schipper 1989, 101).[3]

Earnings management can be obtained by using accounting methods and estimates (i.e., an accrual-based manipulation) or by undertaking transactions that make reported income closer to some target number than it would otherwise be, rather than maximize the firm's discounted expected cash flows (i.e., a real manipulation).[4] This study is an empirical examination of a form of "real" earnings manipulation: the *timing* of income recognition from disposal of assets.

---

[3] For an analysis of recent earnings-management research, see, e.g., Schipper (1989) and Watts and Zimmerman (1990, 138–40).

[4] An example of a real manipulation is the sale of 1,000 acres of land by a Lincoln Savings and Loan subsidiary on March 31, 1987. This transaction enabled Lincoln to recognize an accounting gain of $11 million. Referring to this and some other transactions in his opinion in Lincoln Savings and Loan v. Office of Thrift

Since the timing of asset sales is a manager's choice, and since the principle of acquisition cost underlying the accounting valuation of assets implies that changes in the market value of an asset between acquisition and sale are reported in the period of sale, it follows that the timing of asset sales can be an efficient instrument to manipulate earnings. This potential is illustrated by the following two examples. First, if a firm holding a portfolio of marketable securities as a long-term investment intends to sell an investment, management could typically choose to sell either a particular investment that has unrealized holding gains, thereby increasing reported earnings, or an investment that has unrealized holding losses, thereby decreasing reported earnings. The timing of asset sales may thus depend on management's desire to increase or decrease reported earnings.[5] Second, if a decision to sell an asset has been made and a buyer has been identified close to year-end, management can report the transaction results in the current year by rushing the finalization of the transaction or in the next year by delaying finalization.

Prior studies have explored real earnings manipulation in the context of early debt retirements (see, e.g., Hand 1989; Hand et al. 1990). However, managers are more likely to manipulate earnings through asset sales because: (1) accounting income from asset sales affects income from continuing operations, whereas income from early debt retirements is reported as an extraordinary item; (2) investors employ income from continuing operations, rather than net income, in their decision making (for a discussion of this point, see, e.g., Ronen and Sadan 1981, 106–13); and (3) managers manipulate earnings to affect investors' perceptions (see, e.g., Dye 1988; Trueman and Titman 1988). This implies that the power of the current study's tests is increased, which is particularly desirable in the context of earnings manipulation because a number of studies have failed to detect it and have (partially) attributed the failure to low-power tests (see, e.g., DeAngelo 1986; Liberty and Zimmerman 1986). However, as in prior research, failure to document earnings manipulation by asset sales cannot be evidence of nonmanipulative behavior; earnings can be manipulated with other instruments.

II. Formulation of Hypotheses

Earnings-Smoothing Hypothesis

The earnings-smoothing hypothesis suggests that earnings are manipulated to reduce fluctuations around some level considered normal for the firm. A number of causes may underlie smoothing behavior. Barnea et al. (1975) argue that smoothing is a vehicle for management to convey its earnings expectations within generally accepted accounting principles (GAAP), which do not permit making direct forecasts. Hand (1989, 597) agrees that "[o]ccasional smoothing may be an effective way for a firm to

---

Supervision, Judge Stanley Sporkin states, "The main purpose of certain of these transactions was to create paper profits in order to 'window dress' the financial statements of Lincoln . . ." (*The Wall Street Journal* 1990a).

[5] External validation for this example is found in the financial press. According to *The Wall Street Journal* (1990c), "Critics say that under existing practice, financial institutions can enhance their financial profile by selling securities in which they have a profit while continuing to carry at cost a different batch of the same securities in which they have a paper loss."

provide the stock market with information as to the degree of future persistence of current earnings." He also proposes another possible reason for earnings-smoothing, arguing that it produces an earnings per share (*EPS*) number closer to market expectations. A somewhat similar reason is reported by the financial press. According to *Fortune* (1989, 196), "CEOs know that investors hate surprises, so they try to keep net income trending up a nice straight slope. . . ." Weisbach (1988) finds that, after controlling for stock price performance, CEO turnover is higher when accounting earnings are below the previous year's, suggesting that managers may smooth annual earnings to save their positions. In this context, smoothing takes the form of either reducing earnings manipulatively when they are exceptionally (temporarily) high (to avoid an excessive goal in the following year) or inflating them manipulatively when they are lower than the previous year's earnings. Trueman and Titman (1988) propose that income smoothing is beneficial to firms because it dampens the variance of observed earnings and thereby reduces the firm's cost of borrowing.[6]

However, these explanations have been criticized as relying on unrealistic economic and behavioral assumptions. Watts (1982) argues, for example, that earnings manipulations are pointless because market participants are sophisticated and are able to undo such manipulations. Given this controversy about the motivation for earnings-smoothing, and in line with prior empirical research in the area (see, e.g., Hand 1989; Hand et al. 1990), I focus on whether the data are consistent with earnings-smoothing behavior, rather than on why managers engage in this activity.

To test earnings-smoothing, the level around which income is smoothed must be specified. As in prior empirical research on earnings smoothing (see, e.g., Archibald 1967; White 1970), I assume that the previous year's *EPS* is the target.[7] White (1970, 263–64), for example, advocates the use of this definition for two reasons: (1) it is relatively simple; and (2) it appears more realistic than other definitions, which require managers to attain constant annual growth rates. In the context of the timing of asset sales, these considerations provide the following hypothesis (stated in the alternative form):

H_1: There is a negative correlation between income from asset sales and earnings changes (exclusive of asset sale effects).

Debt-Equity Hypothesis

Corporate lending contracts typically contain accounting-based covenants that alleviate the bondholder-stockholder conflict within the borrowing firm and thereby increase firm value (see, e.g., Smith 1993). Debt covenants are divided into affirmative covenants and negative covenants. Affirmative covenants require borrowing firms to maintain specified levels of accounting-based ratios, e.g., minimum working capital and interest coverage. Negative covenants restrict the financing and investing activities of borrowing firms (e.g., dividend payments and issuance of new debt) unless conditions specified in terms of accounting numbers are satisfied. A covenant that becomes binding imposes costs on a firm, either in renegotiating the bond issue with its purchasers to remove the covenant or in restricting its opportunity set.[8] Since it is costly to violate debt covenants, and since covenants contain accounting-based constraints that are frequently defined in terms of earnings (e.g., as levels of unrestricted retained earnings or as the ratio of net income to interest expense), it follows that managers act to minimize technical violation of accounting-based restrictions in debt agreements by earnings manipulation. To test this prediction, it is standard to assume that leverage is a proxy for the existence and closeness of accounting-based constraints.[9] On the basis of this assumption, it has been proposed that, other things being equal, the larger a firm's debt-equity ratio, the more likely its managers are to shift reported earnings from future periods to a current period and to engage in greater manipulations.[10] In timing asset sales, this debt-equity hypothesis suggests the following testable hypothesis (stated in the alternative form):

---

[6] Notice that if the process that generates the pre-smoothed ("true") earnings is a random walk, smoothing increases rather than decreases the (long-term) variability of earnings. However, if the time-series process of reported earnings (the only numbers observable by researchers) is a random walk, smoothing can reduce (long-term) earnings fluctuations (see, e.g., Watts and Zimmerman 1986, 144–46).

[7] As a sensitivity check for the assumption that the target of earnings-smoothing is the prior year's EPS, the tests were repeated with the target as the previous year's EPS plus a drift (defined as the average earnings change over the previous two years). The results (available from the author upon request) were insensitive to this new definition.

> H_2: There is a positive correlation between income from asset sales and debt-equity ratios.[11]

III. Research Design

The research design has two tasks: to test empirical implications of the earnings-smoothing and debt-equity hypotheses, and to assess the sensitivity of the results to possible specification errors.

Testing of Hypotheses

The empirical investigation begins by testing the two hypotheses separately. Univariate parametric and non-parametric tests (described in detail in the next section) are employed. Although multiple regression tests are more powerful if correctly specified, these univariate tests are an important first step because they rely on fewer assumptions and are thus more robust.[12]

---

[8] The renegotiation costs include the out-of-pocket costs of revising the loan agreement as well as the extra protection or higher interest rate the debt-holder may demand as a compensation for the renegotiation. For public debt, the renegotiation is done indirectly through the trustee, who typically must obtain agreement from at least two-thirds of the debt-holders prior to removing the covenant; if the debt is private, renegotiation is directly with the debt-holders. Also, Smith and Warner (1979) analyzed a sample of bond covenants and reported that 91 percent restricted the issuance of new debt and 23 percent restricted dividend payments.

[9] This choice of leverage as a proxy has been theoretically justified (see, e.g., Gavish and Kalay 1983), empirically confirmed (see, e.g., Duke and Hunt 1990; Press and Weintrop 1990), and is common in prior research (for a partial list see Press and Weintrop 1990, fn. 3). However, as pointed out by Press and Weintrop (1990, 65), if leverage measures closeness with error, the power of tests that use it as a proxy is weakened.

[10] For a detailed discussion of this debt-equity hypothesis, see, e.g., Watts and Zimmerman (1986, chap. 9).

[11] The debt-equity ratio is defined as the ratio of the book value of long-term debt to the book value of owners' equity at the beginning of the event year. As a check for the sensitivity of the results relative to this particular definition, the tests for the debt-equity hypothesis (described later) were repeated with the debt-equity ratio defined in two other ways: (1) as the ratio of book value of long-term debt to the market value of common stock at the beginning of the event year, and (2) as the ratio of the book value of long-term debt to the book value of owners' equity before income from asset sales at the end of the event year. The results for all definitions were indistinguishable, so only those for the first are reported.

[12] For example, univariate tests do not require linearity and are not sensitive to outliers of the independent variables.

Next, the predictions of the two hypotheses are tested simultaneously in a multiple regression, which may be stated formally as follows:

$$ASSIN_i = a_0 + a_1 \delta EPS_i + a_2 DETEQ_i + \epsilon_i. \qquad (1)$$

$$H_{0,1}: a_1 \geq 0, \quad H_{0,2}: a_2 \leq 0,$$
$$H_{a,1}: a_1 < 0, \quad H_{a,2}: a_2 > 0,$$

where $ASSIN_i$ is the income from asset sales per share of the ith firm-year observation for the event year, deflated by the stock price at the beginning of the year; δEPS_i is the change in pre-tax annual ordinary income per share, exclusive of income from asset sales, of the ith firm-year observation for the event year, scaled by the stock price at the beginning of the year; $DETEQ_i$ is the ratio of the book value of long-term debt to the book value of owners' equity of the ith firm-year observation at the beginning of the event year; and ϵ_i is the $ASSIN_i$ residual.

Sensitivity Analysis

A methodological limitation underlying earnings management research is the possible existence of a correlated omitted-variables problem, which results in biased parameter estimates in equation (1); that is, nonzero estimates may be observed even in the absence of earnings manipulation.[13] To determine whether this correlated omitted-variables problem is serious in the current context, the sensitivity of the results of equation (1) is evaluated relative to asset sale levels (i.e., cash generated by asset sales), bonus plans, and possible tax effects.

Asset Sales. Accounting income from asset sales is the realization of previously unrealized holding income that has no cash-flow implications for the firm (possible tax effects aside). Therefore, evidence of earnings-smoothing and debt-equity effects, even after controlling for (i.e., keeping constant) the level of asset sales, decreases the likelihood that the inferences are attributable to misspecified statistical tests.[14]

Introducing asset sales as an additional explanatory variable into equation (1) requires that the relation between income from asset sales and asset sales be monotonic. However, income from asset sales is increasing in asset sales when income is positive and decreasing in asset sales when income is negative. To make the relation monotonic, when income from asset sales is negative, asset sales enter the regression with negative signs.[15]

---

[13] For a formal statistical analysis of this correlated omitted-variables problem, see, e.g., Maddala (1977, 155–57) and Greene (1990, 259–61). For a discussion of this problem in the context of earnings manipulation, see, e.g., Ronen and Sadan (1981, 34–36) and McNichols and Wilson (1988, 4–7). To illustrate this point, consider, for example, a firm whose deteriorating performance leads to a liquidity crunch, which in turn leads to a high level of asset sales to satisfy cash needs. Further, assume that asset sales and accounting income from asset sales are positively correlated (because of inflation). These conditions lead to predictions similar to those of the earnings-smoothing hypothesis, even if asset sales result solely from managers' optimal investing and financing decisions (i.e., without earnings manipulation).

[14] To see this, note that a certain level of earnings management is attained by two, not mutually exclusive, techniques. First, managers can sell assets that in turn change accounting income (technique one). In this scenario, however, an identification problem arises; it is difficult to determine empirically whether the sale represents optimal investing, financing, and operating strategies or earnings manipulation. The second way (technique two) to attain a certain level of earnings management is, for a given asset sale level (determined by a firm's optimal investing, financing, and operating strategies), to sell assets on the basis of the *accounting* income they generate (as in the first example in section I). Identification problems do not arise with this second technique, since real effects (i.e., cash generated by asset sales) are controlled. It seems reasonable to assume that both techniques are employed in managing earnings. Thus, if the findings show that reported income from asset sales and firm performance (debt-equity ratios) are negatively (positively) associated even when the level of asset sales is kept constant, then these omitted variables present no serious problem.

Bonus Plans. The bonus-plan hypothesis assumes that managers maximize their compensation through earnings manipulation and suggests that managers reduce earnings when earnings realizations are outside the lower or upper bounds of the plan and enhance earnings otherwise (see, e.g., Healy 1985). The earnings-smoothing effect and the bonus-plan effect thus partially overlap, and one must control for the latter in the cross-sectional analysis.[16]

Since the relation between earnings and the incentives to manipulate earnings is discontinuous around the lower bound of the bonus plan, translating this relation into a linear model requires the introduction of two variables. The first equals the upper bound of the plan less pre-tax income of the ith firm-year observation for the event year (scaled by its market value of common equity at the beginning of the year) if the pre-tax income exceeds the lower bound, and 0 otherwise. This variable measures the incentives for managers to manipulate earnings when the realized pre-tax income exceeds the lower bound of the bonus plan, since bonuses in these situations are maximized at the point where reported earnings are equal to the upper bound. The second variable is a dummy variable set to 1 if the lower bound of the plan exceeds pre-tax income, and 0 otherwise. This second variable controls for situations in which actual pre-tax earnings are below the lower bound of the bonus plan. A dummy (rather than continuous) variable is used because there is no explicit reported earnings target; the incentive of a bonus-maximizing manager is to report earnings as low as possible.[17]

Controlling for the bonus-plan effect requires knowledge of the bounds of the plan. Since these bounds are available for only a small subset of the sample firms, and in line with an argument in prior research that the bonus-plan hypothesis can be generalized to firms without explicit plans (see, e.g., McNichols and Wilson 1988, 3–4), I assume that the lower and upper bounds of the plan are 10 and 20 percent, respectively, of the firm's net worth at the beginning of the year, and that the bonus is based on pre-tax income.[18]

Tax Effects. The third possible source of correlated omitted variables considered is tax effects. Assuming that the income-tax rate is fixed across time and income, the present value of a future deduction is less than the value of an immediate deduction of the same magnitude, due to positive interest rates and a probability greater than zero of future bankruptcy. Under certain circumstances, a company can thus time asset sales to reduce its current tax liability and thereby save taxes. Such tax savings can be obtained when there is no carry-forward tax loss, taxable income is positive, and a loss from sales of depreciable assets for income-tax reporting is recognized.[19] In other situations, income from asset sales is unlikely to reduce a firm's current tax liability.

---

[15] Alternatively, a dummy variable technique can be used to allow the slopes of income from asset sales (with respect to asset sales) to vary across negative and positive incomes. Empirically, these two alternative models yielded similar results.

[16] For example, if earnings before asset sale income are up from last year and also above the upper bound of the bonus plan, managers might sell assets with unrealized holding losses to bring the reported income closer to either last year's earnings (the earnings-smoothing effect) or to the upper bound of the bonus plan (the bonus-plan effect).

[17] This variable also (partially) captures the "big bath" phenomenon: if a firm has a loss, managers increase the loss by including future write-offs to improve future earnings. However, this is not a problem because no inferences are drawn about the validity of the bonus-plan hypothesis. The analysis in this section is intended only to verify that the results for the debt-equity and earnings-smoothing hypotheses are robust to a possible omitted-variables problem.

[18] Data presented in prior research (see, e.g., Fox 1979, table 11) suggest that these are typical bounds. As a (partial) check for the sensitivity of the results to these specific bounds, the regression analysis was replicated by using lower (upper) bounds of 5 percent (10 percent) and 20 percent (30 percent). Although the estimates for the parameters of the bonus variables were sensitive to the values of the bounds, the estimates for the parameters of the earnings-smoothing and debt-equity variables were not.

Since tax savings considerations sometimes influence the timing of asset sales, a proxy for possible tax-saving effects is added to the cross-sectional analysis to control for these effects. The proxy is a dummy variable set to 1 if a loss from asset sales is reported and otherwise taxable income is positive, and set to 0 otherwise.[20]

IV. Data Description

The data for this investigation are from the COMPUSTAT expanded annual industrial file with OTC firms, the COMPUSTAT quarterly industrial file, and the CRSP NYSE/ASE and NASDAQ daily returns files. The sample covers the three-year period 1987–1989 because data on income from disposal of assets are not available on COMPUSTAT prior to 1987. This restriction limits the ability to generalize the results to other time periods.

The sample contains firm-year observations with income from asset sales reported as ordinary income during the sample period.[21] In addition, three criteria were imposed in order to restrict the empirical analysis to observations for which data are readily available in a machine-readable form, to remove outliers, which is common in the literature (see, e.g., Beaver et al. 1987, table 1), and to ensure that the results are not driven by a small number of firms that frequently sell assets.[22] These three conditions represent trade-offs between cost-effectiveness, increased power and validity of the tests, and the generality of the results.

The final sample consists of 653 firm-year observations.[23] Observations reporting

---

[19] The U.S. Internal Revenue Code (IRC) permits the deduction of a net capital loss in the year it is sustained only if it results from sales of depreciable assets (see IRC Sections 1211(a) and 1231; IRC Reg. Section 1.1211-1).

[20] Notice that it is possible that taxable gains are recognized even when accounting losses from sales of depreciable assets are reported in financial statements, because most corporations use accelerated depreciation for tax purposes and straight-line depreciation for financial reporting (see, e.g., Horngren and Sundem 1990, 327). This situation, as well as the fact that income-tax rates vary across time and income, implies that the proxy for tax effects is noisy. While under some circumstances this may bias the statistical tests, sensitivity analysis (not reported here) shows that the results that follow are robust relative to the choice of a proxy for tax effects.

[21] In its Opinion No. 30 (APB 1973), the Accounting Principles Board stipulates criteria for classification of events and transactions as extraordinary items. To qualify, they must be both unusual and nonrecurring. The Board also specified several events and transactions that do not constitute extraordinary items, such as gains or losses from sale or abandonment of property, plant, or equipment used in the business.

[22] The three criteria are: (1) data on the following variables were available on the COMPUSTAT expanded annual industrial file with OTC firms: pre-tax annual income (item 170), long-term debt (item 9), (positive) book value of stockholders' equity (item 216), stock price at fiscal year-end (item 199), (ordinary) income from asset sales (item 213), sales of investments (item 109), sales of long-lived assets (item 107), common shares used to calculate EPS (primary) (item 54), adjustment factors for stock dividends and splits (item 27); (2) sales of long-lived assets, sales of investments, and the change in pre-tax annual earnings (in absolute value), all deflated by the market value of common equity at the beginning of the year, did not exceed 50 percent (30 percent for the last ratio); (3) a company could be represented in the sample by no more than one firm-year observation. When two or more firm-year observations for the same company were available, only the first occurrence (arbitrarily chosen) was included in the sample.

[23] The selection procedure yields 658 observations, but four observations with unusually high debt-equity ratios and one observation with unusually high sales of investments (10–40 s.e. from the mean) were considered outliers. The inclusion of these five observations in the sample did not qualitatively affect any of the conclusions. See also footnote 27.

Table 1
Descriptive Statistics for Sample and COMPUSTAT Firms

	Sample			*Other COMPUSTAT Firms*			
Statistic	*Q1*	*Median*	*Q3*	*Q1*	*Median*	*Q3*	*z-statistic*[a] *p-value*
Market value of common equity ($million)	30.152	81.808	267.258	38.599	136.261	626.550	0.00
Annual sales ($million)	46.500	125.813	370.581	59.966	229.684	988.311	0.00
Total assets ($million)	45.275	111.779	309.264	66.836	264.810	1655.200	0.00
Current ratio[b]	1.533	2.144	3.004	1.234	1.820	2.701	0.00

Note: The comparison is for FY 1988 (arbitrarily chosen). The number of observations used for computing the statistics ranges with data availability from 625 to 653 for the sample and from 2,068 to 2,627 for COMPUSTAT firms not in the sample.

[a] The z-statistic tests the null that the median of the sample firms equals that of the other COMPUSTAT firms by using the two-sample distribution-free median test; observed significance levels for two-tailed tests are reported.

[b] The ratio of current assets to current liabilities.

gains (losses) from asset sales are spread over 56 (51) different two-digit standard industrial classification (SIC) industries. Thus, a broad cross-section of firms is represented in the final sample. Grouped by two-digit SIC industry codes, the 397 firms reporting gains have a median group size of four; the median group size for the 256 firms reporting losses is three. There is no evidence of clustering by industry in the sample.

A comparison of the first, second, and third quartiles of four descriptive statistics is shown in table 1 for the sample and for all COMPUSTAT firms not in the sample (FY 1988 is arbitrarily chosen). These four statistics include three firm-size measures, market value of common equity, total annual sales, and total assets, and a short-term liquidity measure, the current ratio (i.e., the ratio of current assets to current liabilities). Firm size was selected due to its importance in explaining cross-sectional differences with respect to such characteristics as firm performance and risk, and the current ratio was selected as a measure for the short-term soundness of the firm. Although the sample clearly spans a wide range of firm sizes, sample firms are substantially smaller than other COMPUSTAT firms. However, the current ratios of the sample firms are larger than those of the other COMPUSTAT firms; the sample firms thus appear financially healthy.

V. Empirical Results

Univariate Tests for the Earnings-Smoothing Hypothesis

Univariate tests of the earnings-smoothing hypothesis are in table 2, which shows summary statistics on one subsample containing firms with a positive earnings change before asset-sale income ($\delta EPS_t > 0$) in the event year and on another subsample with firms that experience a negative earnings change ($\delta EPS_t < 0$). Firms with a positive earnings change appear larger and, to some extent, less leveraged than those with a negative earnings change. The current ratios of the two subsamples are almost iden-

Table 2
Summary Statistics for Firms that Experience Positive and Negative Earnings Changes for the Event Year

	Negative δEPS (n=294)		*Positive δEPS* (n=359)		*z-statistic*[b] *p-value*	
Statistic[a]	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*
δEPS	−4.63%	−7.48%	4.24%	5.66%	—	—
Market value of common equity ($million)	84.124	281.378	124.865	549.398	0.12	0.00
Common-stock beta	0.81	0.92	1.02	1.05	0.00	0.01
Debt-equity ratio	40.81%	64.71%	37.81%	53.30%	0.72	0.13
Current ratio	2.10	2.71	2.15	2.53	0.54	0.33
Sale of investments divided by market value of common equity	0.00%	0.35%	0.00%	0.41%	0.05	0.58
Sale of fixed assets divided by market value of common equity	0.52%	1.22%	0.50%	0.78%	0.72	0.00
Total sales of assets divided by value of common equity	1.49%	1.58%	1.32%	1.19%	0.04	0.03
Income from asset sales scaled by market value of common equity	0.09%	1.23%	0.03%	0.53%	0.03	0.02

Note: The stock prices and market values of equity used as deflators are as of the beginning of the event year.

[a] δEPS is the change in pre-tax annual ordinary income, exclusive of income from asset sales, deflated by beginning-of-the-year stock price for the event year. Common-stock beta (beta parameter of the market model) of a firm is estimated by using daily stock returns and equally weighted market returns for 100 days commencing at the beginning of the event year. Only firms with a minimum of 50 daily returns are included. Only 322 (275) firms experiencing positive (negative) δEPS had the necessary data on the CRSP NYSE/AMEX and NASDAQ daily returns files for estimation of their common-stock beta. Debt-equity ratio is the ratio of the book value of long-term debt to the book value of owners' equity at the beginning of the event year. Current ratio is the ratio of current assets to current liabilities at the beginning of the event year.

[b] The z-statistic tests the null that the median (mean) of the negative δEPS firms equals that of the positive δEPS firms by using the two-sample distribution-free median (student-t) test. For the first eight statistics, observed significance levels for two-tailed tests are reported; for the last statistic, one-tailed tests are reported.

tical. However, firms that experience positive changes seem to be riskier (the differences between the two subsamples' median and mean common-stock betas are significant at 1 percent levels for two-tailed tests).[24] Comparisons of sales of investments and long-lived assets are inconclusive. For both samples, however, the medians and means of sales of investments and long-lived assets are generally less than 1 percent, which suggests that these sales do not represent a major disposal of assets.

The earnings-smoothing hypothesis implies that firms that exhibit a negative earnings change before asset sale income $(\delta EPS_i<0)$ should report higher income from asset sales than firms that experience a positive change $(\delta EPS_i>0)$. The findings

---

[24] The common-stock beta (beta parameter of the market model) of a firm is estimated by using daily stock returns and equally weighted market returns for 100 days commencing at the beginning of the event year. Only firms with a minimum of 50 daily returns are included.

Table 3
Asset Sales as a Percentage of Annual Asset Sales by Fiscal Quarter

	Long-Lived Assets				Investments			
Fiscal Quarter	Median	z-statistic p-value	Mean	t-statistic p-value	Median	z-statistic p-value	Mean	t-statistic p-value
1	0.6	0.000	5.3	0.000	1.9	0.000	11.2	0.000
2	11.4	0.000	12.7	0.000	12.9	0.000	17.3	0.000
3	25.0	0.000	21.8	0.000	26.3	0.004	27.4	0.000
4	52.6	—	60.2	—	38.1	—	44.1	—

Notes: Asset sales are measured in dollar amounts. The z-statistic (t-statistic) tests the hypothesis that the median (mean) asset sales for the fourth quarter exceeds that for quarter *i* (i=1,2,3). Observed significance levels (p-values) for one-tailed tests are reported.

presented in table 2 are consistent with this prediction; in particular, the median and mean income from asset sales (scaled by market value of common equity) for the $\delta EPS<0$ firms are 0.09 and 1.23 percent, respectively, compared to 0.03 and 0.53 percent for the $\delta EPS_t>0$ firms (the differences are significant at the 5 percent level for one-tailed tests).[25]

Finally, the timing of asset sales by fiscal quarter is examined. Prior research (see, e.g., Hand et al. 1990, 78) has shown that real earnings-smoothing activities (through in-substance defeasances) cluster in the fourth fiscal quarter. The evidence in table 3 on asset sales (measured in dollar amounts) by fiscal quarter reveals a similar pattern. The percentage of assets sold by the sample firms in each fiscal quarter increased monotonically during the fiscal year. In addition, the difference between the median (mean) percentage of asset sales for the fourth quarter and the median (mean) for any of the other three fiscal quarters is statistically significant at the 1 percent level for one-tailed tests. Furthermore, the sale of long-lived assets in the fourth fiscal quarter exceeds that in the three other fiscal quarters combined.

Univariate Tests for the Debt-Equity Hypothesis

The empirical investigation of the debt-equity hypothesis begins by partitioning the data into high- and low-leverage firms on the basis of the median debt-equity ratio of the sample. The summary statistics in table 4 indicate that the samples are similar with respect to changes in pre-tax EPS (exclusive of asset sales), firm size, and common-stock betas (the differences between the medians and means for all three characteristics are insignificant at 10 percent levels for two-tailed tests). These findings increase confidence that the following results are not attributable to correlated omitted variables.

---

[25] The median and mean incomes from asset sales for both subsamples are positive. The earnings-smoothing hypothesis does not necessarily imply that income from asset sales of companies above the target EPS, which tend to smooth earnings downward, is negative, merely that it is below income from asset sales in the absence of earnings manipulations. However, the positive mean and median incomes from asset sales of positive δEPS firms open the possibility that there is "single-sided" income smoothing going on; i.e., earnings are smoothed primarily when the earnings change is negative. This possibility is examined later.

Table 4
Summary Statistics for High- and Low-Leverage Firms

	High Leverage (n=327)		*Low Leverage* (n=326)		*z-statistic*[b] *p-value*	
Statistic[a]	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*
Debt-equity ratio	76.27%	106.04%	11.20%	12.90%	—	—
δEPS	0.33%	−0.42%	0.72%	−0.08%	0.41	0.64
Market value of common equity ($million)	72.282	415.564	80.748	442.922	0.33	0.78
Common-stock beta	0.93	0.99	0.92	0.99	0.83	0.93
Current ratio	1.94	2.26	2.32	2.95	0.01	0.01
Sale of investments divided by market value of common equity	0.00%	0.39%	0.00%	0.37%	0.21	0.84
Sale of fixed assets divided by market value of common equity	0.50%	1.28%	0.17%	0.67%	0.00	0.00
Total sales of assets divided by market value of common equity	0.78%	1.68%	0.35%	1.05%	0.00	0.00
Income from asset sales scaled by market value of common equity	0.14%	1.28%	0.02%	0.41%	0.00	0.00

Note: The stock prices and market values of equity used as deflators are as of the beginning of the event year.

[a] High- and low-leverage firms are those with debt-equity ratios larger (smaller) than the median debt-equity ratio of the full sample. Only 304 and 293 firms of the high- and low-leverage firms, respectively, had the necessary data on the CRSP NYSE/AMEX and NASDAQ daily returns files for estimation of their common-stock beta. For definitions of debt-equity ratio, δEPS, common-stock beta, and current ratio, see table 2.

[b] The z-statistic tests the null that the median (mean) of low-leverage firms equals that of the high-leverage firms by using the two-sample distribution-free median (student-*t*) test. For the first eight statistics, observed significance levels for two-tailed tests are reported; for the last statistic, one-tailed tests are reported.

The results also indicate that high-leverage firms have significantly lower current ratios than low-leverage firms. In addition, high-leverage firms sell more long-lived assets than do low-leverage firms. The median and mean sales of long-lived assets as a percentage of the market value of common equity of high-leverage firms significantly exceed (at a 1 percent level for two-tailed tests) those of low-leverage firms. The difference between the two samples with respect to sales of investments is, however, insignificant at conventional levels. Also, the medians and means of total asset sales (deflated by the market value of common equity) for both subsamples are around 1 percent, an indication that these sales do not represent unusual selling activities (e.g., a major restructuring).

The debt-equity hypothesis implies that income from asset sales by higher-leverage firms exceeds that for lower-leverage firms; this is consistent with the results reported in table 4.

Multiple Regression Tests

In this section, equation (1) is used to test simultaneously the predictions of the earnings-smoothing hypothesis (that a_1 is negative) and the debt-equity hypothesis (that

Table 5
Multiple Regression Tests of Debt-Equity and Earnings-Smoothing Hypotheses

Model: $ASSIN_i = a_0 + a_1 \delta EPS_i + a_2 DETEQ_i + \epsilon_i$.[a]

Sample	a_0	a_1	a_2	*Model p-value*[b]	*Adjusted* R^2
Expected sign:		−	+		
Full sample ($n=653$)	0.003 (0.06)	−0.071 (0.00)	0.008 (0.00)	0.000	0.068
Postive δEPS_i ($n=359$)	0.004 (0.04)	−0.027 (0.09)	0.004 (0.01)	0.027	0.015
Negative δEPS_i ($n=294$)	−0.004 (0.26)	−0.137 (0.00)	0.010 (0.00)	0.000	0.093

Note: Under the alternative hypothesis of earnings manipulation, $a_1 < 0$, $a_2 > 0$. Observed significance levels (i.e., p-values) for one-tailed tests are in parentheses (two-tailed tests for the intercept).

[a] $ASSIN_i$ is income from asset sales of the *i*th firm for the event year deflated by the stock price at the beginning of that year; $DETEQ_i$ is the long-term debt-to-equity ratio of the *i*th firm at the beginning of the event year; δEPS_i is the change in pre-tax annual ordinary income, exclusive of income from asset sales, for the event year, scaled by beginning-of-the-year stock price.

[b] Tests the hypothesis that all model parameters (excluding the intercept) are zero.

a_2 is positive). This multiple regression analysis provides a method for determining the incremental effect of each variable and the overall significance of the variables.

Summary results for the full sample and the two subsamples formed on the basis of the sign of δEPS_i are reported in table 5. For the full sample, the estimate for the coefficient of the change in annual income exclusive of income from asset sales (δEPS_i), a_1, is negative and significant (at the 1 percent level for one-tailed tests), and the estimate for the coefficient of the debt-to-equity ratio ($DETEQ_i$), a_2, is positive and significant (at the 1 percent level for one-tailed tests). The results also show that the test that the earnings-smoothing and the debt-equity effects are jointly zero can be rejected with a high level of confidence.

The results for the two subsamples demonstrate two phenomena. First, the earnings-smoothing effect is documented in both subsamples, which suggests that both positive and negative earnings changes may be subject to smoothing. This finding suggests that the conjecture in research on distressed companies that these firms employ single-sided income smoothing cannot be generalized to my sample of financially healthy firms (see also fn. 25). Second, the slope estimates for the subsample of firms with negative δEPS_i exceed (in absolute values) the corresponding estimates for the subsample of firms with positive δEPS_i (the differences are significant at the 10 percent level for one-tailed tests). This implies that the earnings-smoothing and the debt-equity effects are more pronounced for negative δEPS_i firms.

Sensitivity Analysis

To assess the sensitivity of the findings documented above to possible specification errors (i.e., correlated omitted variables), the effects of asset sale levels, management bonus plans, and tax considerations on the timing of reporting income from asset sales are considered, by estimating the following model:

Table 6
Sensitivity Analysis
(n = 653)

Model: $ASSIN_i = a_0 + a_1 \delta EPS_i + a_2 DETEQ_i + a_3 SALEP_i + a_4 SALEIN_i + a_5 BONUS1_i + a_6 BONUS2_i + a_7 TAX_i + \epsilon_i$.[a]

Expected sign:	–	+	+	+	+	–	–		
a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	Model p-value[b]	Adjusted R^2
0.007 (0.00)	−0.0722 (0.00)	0.007 (0.00)	0.242 (0.00)	0.418 (0.00)	−0.021 (0.16)	−0.005 (0.06)	−0.011 (0.00)	0.00	0.160

[a] $ASSIN_i$ is income from asset sales of the ith firm for the event year deflated by the stock price at the beginning of that year; $DETEQ_i$ is the long-term debt-to-equity ratio of the ith firm at the beginning of the event year; δEPS_i is the change in pre-tax annual ordinary income, exclusive of income from asset sales, for the event year, scaled by beginning-of-the-year stock price; $BONUS1_i$ is the upper bound of the bonus plan less pre-tax income of the ith firm-year observation for the event year scaled by the market value of common equity at the beginning of the year if the pre-tax income exceeds the lower bound, 0 otherwise; $BONUS2_i$ is a dummy variable set to 1 if the lower bound of the plan exceeds pre-tax income, set to 0 otherwise; $SALEP_i$ ($SALEIN_i$) is sales of long-lived assets (investments) of the ith firm in the event year scaled by the market value of common equity at the beginning of the event year (sales take on negative signs if $ASSIN_i$ is negative); and TAX_i is a dummy variable set to 1 if a loss from asset sales is reported and if taxable income exclusive of this loss from asset sales is positive, set to 0 otherwise.

[b] Tests the hypothesis that all model parameters (excluding the intercept) are zero.

$$ASSIN_i = a_0 + a_1 \delta EPS_i + a_2 DETEQ_i + a_3 SALEP_i + a_4 SALEIN_i + a_5 BONUS1_i + a_6 BONUS2_i + a_7 TAX_i + \epsilon_i, \qquad (2)$$

where $ASSIN_i$, δEPS_i, $DETEQ_i$, and ϵ_i are as in equation (1); $SALEP_i$ ($SALEIN_i$) is the sales of long-lived assets (investments) of the ith firm in the event year, scaled by the market value of common equity at the beginning of the event year (sales take on negative signs when income from asset sales is negative); $BONUS1_i$ is the upper bound of the plan less pre-tax income of the ith firm-year observation for the event year scaled by the market value of common equity at the beginning of the year if the pre-tax income exceeds the lower bound, 0 otherwise; $BONUS2_i$ is a dummy variable set to 1 if the lower bound of the plan exceeds pre-tax income and set to 0 otherwise; and TAX_i is a dummy variable set to 1 if a loss from asset sales is reported and taxable income, exclusive of this loss from asset sales, is positive and set to 0 otherwise.[26]

The earnings-smoothing hypothesis predicts that the coefficient estimate of δEPS_i, a_1, is negative, and the debt-equity hypothesis predicts that the coefficient estimate of $DETEQ_i$, a_2, is positive. The results in table 6 are consistent with these predictions. In addition, the coefficient estimates of δEPS_i and $DETEQ_i$ for equations (1) and (2) appear quite similar (cf. tables 5 and 6). These results increase the confidence that possible specification errors do not present a serious problem.

With respect to the parameter estimates of the other variables of equation (2), the results are significant with the expected signs for the effects of asset sale levels and tax considerations and mixed for the effect of bonus plans. Specifically, a_3 and a_4, the estimates for the coefficients of the variables $SALEP_t$ and $SALEIN_t$ that control for the level of asset sales, are highly significant; a_5, the estimate for the coefficient of the bonus-plan variable $BONUS1_t$, is insignificant at conventional levels, where a_6, the estimate for the coefficient of the bonus plan variable $BONUS2_t$, is marginally significant. The coefficient of TAX_t, a_7, is negative and highly significant, suggesting that tax considerations play a role in timing asset sales. Other specification tests (not reported) further corroborate that the findings are robust to financial soundness, types of asset sales, and possible tax effects.[27]

---

[26] Taxable income is estimated on the basis of the annual accounting tax expense adjusted for all deferred and prepaid income taxes.

VI. Summary

This study explores empirically whether managers manipulate reported earnings through the timing of sales of long-lived assets and investments by taking advantage of the acquisition-cost principle underlying the accounting valuation of assets. Several empirical implications of the earnings-smoothing and debt-equity hypotheses in the context of asset sales were examined. The findings are consistent with the timing of asset sales by managers so that the recognized accounting income from these sales smooths intertemporal earnings changes and mitigates accounting-based restrictions in debt covenants.

In conformance with the earnings-smoothing hypothesis, the findings show that income from asset sales is significantly higher for firms that exhibit decreases in annual earnings (exclusive of asset sales effects) than for firms experiencing increases. In conformance with the debt-equity hypothesis, the evidence indicates that income from asset sales of high debt-equity firms significantly exceeds that of low debt-equity firms.

Results from a cross-sectional OLS analysis that regressed income from asset sales on two explanatory variables (annual earnings changes and debt-equity ratios) indicate that the earnings-smoothing and debt-equity effects are incremental; after controlling for one effect, the other still exists. Finally, the documented earnings-smoothing and debt-equity effects are robust to specifications that consider asset sale levels, bonus plans, possible tax effects, and current ratios. These findings increase the confidence that the results are not attributable to correlated omitted variables.

---

[27] Two diagnostic tests were also performed. One assessed whether outliers drove the regression's results by replicating equation (2) with ranks in place of the continuous independent variables. The results (not reported) were similar to those reported in table 6 with one exception: a_3, the coefficient of $SALEP_t$, became insignificant. The analysis used 658 observations (see fn. 23). Another diagnostic test involved collinearity analysis. Tests based on Belsley et al.'s (1980) approach were employed. The results (not reported) show that all condition numbers are small, ranging between 1 and 4, which indicates that a multicollinearity problem does not exist.

References

Accounting Principles Board. 1973. *Opinion No. 30: Reporting the Results of Operations*. New York: AICPA.

Archibald, T. R. 1967. The return to straight line depreciation: An analysis of a change in accounting methods. *Journal of Accounting Research* 4 (Supplement): 164–80.

Barnea, A., J. Ronen, and S. Sadan. 1975. The implementation of accounting objectives: An application to extraordinary items. *The Accounting Review* 50 (January): 58–68.

Beaver, W. H., R. A. Lambert, and S. G. Ryan. 1987. The information content of security prices: A second look. *Journal of Accounting & Economics* 11 (July): 139–58.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics*. New York: John Wiley & Sons.

DeAngelo, L. E. 1986. Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders. *The Accounting Review* 61 (July): 400–20.

Duke, J. C., and H. G. Hunt III. 1990. An empirical examination of debt covenant restrictions and accounting-related debt proxies. *Journal of Accounting & Economics* 12 (January): 45–63.

Dye, R. A. 1988. Earnings management in an overlapping generations model. *Journal of Accounting Research* 26 (Autumn): 195–235.

Fortune. 1989. Cute tricks on the bottom line. (April 24): 193–200.

Fox, H. 1979. Top executive bonus plans. Conference Board Report No. 760. New York: Conference Board.

Gavish, B., and A. Kalay. 1983. On the assets substitution problem. *Journal of Financial and Quantitative Analysis* 18 (March): 21–30.

Greene, W. H. 1990. *Econometric Analysis*. New York: Macmillan.

Hand, J. R. 1989. Did firms undertake debt-equity swaps for an accounting paper profit or true financial gain? *The Accounting Review* 64 (October): 587–623.

——, P. J. Hughes, and S. E. Sefcik. 1990. Insubstance defeasances: Security price reactions and motivations. *Journal of Accounting & Economics* 13 (May): 47–89.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting & Economics* 7 (April): 85–107.

Horngren, C. T., and G. L. Sundem. 1990. *Introduction to Financial Accounting*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall.

Kieso, D. L., and J. J. Weygandt. 1987. *Intermediate Accounting*, 6th ed. New York: John Wiley & Sons.

Liberty, S., and J. Zimmerman. 1986. Labor union contract negotiations and accounting choices. *The Accounting Review* 61 (October): 692–712.

Maddala, G. S. 1977. *Econometrics*. New York: McGraw-Hill.

McNichols, M., and G. P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

Press, E. G., and J. B. Weintrop. 1990. Accounting-based constraints in public and private debt agreements. *Journal of Accounting & Economics* 12 (January): 65–95.

Ronen, J., and S. Sadan. 1981. *Smoothing Income Numbers: Objectives, Means, and Implications*. Reading, MA: Addison-Wesley.

Schipper, K. 1989. Commentary on earnings management. *Accounting Horizons* 3 (December): 91–102.

Smith, C. W. 1993. A perspective on accounting-based debt covenant violations. *The Accounting Review* 68 (April): 289–303.

——, and J. B. Warner. 1979. On financial contracting: An analysis of bond covenants. *Journal of Financial Economics* 7 (June): 117–61.

The Wall Street Journal. 1985. Small business: Good-looking balance sheets can be assets for companies. September 23.

——. 1990a. At Lincoln Saving: Skullduggery. September 19.

——. 1990b. Accounting group says another panel will handle debt security guidelines. October 31.

——. 1990c. Accounting panel clears new rules for debt issues. September 14.

Trueman, B., and S. Titman. 1988. An explanation for accounting income smoothing. *Journal of Accounting Research* 26 (Supplement): 127–39.

Watts, R. L. 1982. Does it pay to manipulate EPS? *Chase Financial Quarterly* 2 (Spring): 8–17.

——, and J. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice-Hall.

——, and ——. 1990. Positive accounting theory: A ten year perspective. *The Accounting Review* 65 (January): 131–56.

White, G. 1970. Discretionary accounting decisions and income normalization. *Journal of Accounting Research* 8 (Autumn): 260–74.

Weisbach, M. 1988. Outside directors and CEO turnover. *Journal of Financial Economics* 20 (January–March): 431–60.

SHORTER ARTICLES

1. Alex Dontoh and Joshua Ronen
   Information Content of Accounting Announcements

2. Robert E. Verrecchia
   How Do We Assess a Model of Price and Volume?

3. Richard C. Sansing
   Information Acquisition in a Tax Compliance Game

4. Kwang-Hyun Chung, Dimitrios Ghicas, and Victor Pastena
   Lenders' Use of Accounting Information in the Oil and Gas Industry

5. Grace Pownall, Charles Wasley, and Gregory Waymire
   The Stock Price Effects of Alternative Types of Management Earnings Forecasts

6. Stephen P. Baginski, Edward J. Conrad, and John M. Hassell
   The Effects of Management Forecast Precision on Equity Pricing and on the Assessment of Earnings Uncertainty

7. Charles R. Enis
   Earnings-Based Compensation Plans, Performance, and Capital Expenditure Policy in the Motor Carrier Industry

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 857–869

Information Content of Accounting Announcements

Alex Dontoh
Joshua Ronen
New York University

SYNOPSIS AND INTRODUCTION: The efficient market hypothesis (EMH) has spawned a large and impressive body of "event studies." Conditional on the maintained hypothesis of market efficiency, these studies test the "information content" of events such as accounting earnings releases (Fama 1991). However, the notion of "information content" of accounting earnings has not been formally defined, except implicitly—via the way it is calibrated in event studies—as percent price change (risk and size adjusted) associated with earnings releases. In the assumption of belief homogeneity underlying the original formulation of EMH, where no competitive non-null trading is expected in equilibrium, price changes will reflect fully the public announcement (see, e.g., Grossman and Stiglitz 1980). Hence, price changes will, conditional on the endowment and preferences configuration and on prior homogeneously held beliefs, constitute a one-to-one mapping with the belief change induced by the earnings release, and could be adopted as an operational definition of information content (*IC*).

However, admitting the possibility of belief-change heterogeneity into the analysis brings into question the theoretical validity and plausibility of this operational definition. Competitive trading that becomes possible as a result of disagreement would perturb the one-to-one mapping between price change and belief change induced by the earnings release. In any case we no longer have a well-defined single construct reflecting heterogeneous belief changes. Under belief homogeneity the market acts as if it were a single decision-maker whose belief change theoretically reflects and calibrates the value or "content" of information.[1] Once we allow for

---

[1] Hirshleifer (1973) considers information broadly as events that change individuals' beliefs (or subjective probability distributions) over possible states of the world. He distinguishes this view from the notion of information concerned with reducing uncertainty over possible states of the world.

We thank Rowland Atiase, Peter Cheng, Bharat Sarath, Joshua Livnat, Maureen McNichols, Paul Pfleiderer, Catherine Shalen, and workshop participants at Baruch College, New York University, Purdue University, the University of British Columbia, and participants at the 1990 Annual Conference of the American Accounting Association for their comments and suggestions on an earlier version. We also would like to thank an anonymous reviewer for useful suggestions.

Submitted June 1989 and April 1993.
Accepted May 1993.

multiple decision-makers, or equivalently, for belief-heterogeneity, however, this simple definition of *IC* is no longer applicable. How is one to aggregate the belief changes of different traders, and what weights should be attached to such changes in the process of aggregation? Moreover, even if an appropriate operational definition of *IC* were to be discovered under such circumstances, what would be the empirical manifestations?

The derived "information content" measure includes volume reaction, price reaction, and the predisclosure expected belief dispersion. The results imply that neither price reaction nor volume reaction alone will provide a complete characterization of information content; both must be jointly observed along with some aspects of the existing information environment to gauge information content fully.

Key Words: *Public disclosure, Information content, Heterogeneous inferences.*

I. Development of an Information Content Measure

IN this paper, we address these questions and develop an operational definition of information content in a regime of heterogeneous beliefs that is derived from a three-date, two-trading-rounds, noisy rational expectation equilibrium model of trading and prices. After the first round of trading induced by heterogeneous private signals, traders interpret a public accounting announcement heterogeneously as they adjust their portfolios in the second round. The *IC* definition we develop includes volume reaction, price reaction, and the predisclosure expected belief dispersion; decreases in the dispersion in individual inferences from the public disclosure; and increases in the precision of the disclosure. Separately, trading reaction increases in both the precision of the public signal and the predisclosure expected dispersion, but it is nonmonotonic with respect to the dispersion in traders' interpretations of the public disclosure (first increasing then decreasing). The results imply that neither price reaction nor volume reaction alone will provide a complete characterization of *IC*; *both must be jointly observed* along with some aspects of the existing information environment to gauge information content.

While existing price and trading equilibrium models (see Karpoff 1987) investigate the impact of belief revision heterogeneity on prices and trading volume, no researchers have addressed the implications of these models for the appropriateness of gauging content in event studies by price changes (abnormal returns), nor have researchers developed a measure of content that incorporates the effects of belief heterogeneity. None of these studies has addressed major institutional features of the accounting environment within which earnings are announced: the flexibility accorded management in accounting choices, the "distortion" such flexibility could create in reported earnings, and the concomitant divergence in the interpretations of the earning announcement.[2] In light of the importance of earnings quality (recent empirical works investigate the impact of some quality characteristics on Earning Response Coefficients [Bernard 1992]), the FASB recognizes the presence of this "distortion" and the likelihood of resulting divergent interpretations among traders. The FASB imposes verifiability as one of the recognition criteria—where verifiability is defined in terms of the degree of consensus on the measurement of an accounting item.[3]

---

[2] While reported earnings result from "strategic" managerial accounting choices, traders, who lack the information to predict these strategies, would view the outcome of the "strategic" choices as distortion noise.

This institutional feature results from an equilibrium which takes into account managerial incentives to disclose information strategically. Regulatory bodies, aware of such incentives, have made decisions that trade off uniform accounting methods that generate comparability and flexibility among accounting alternatives that allows managers to convey private information in accounting reports which are subject to verification and attestation. Generally accepted accounting principles (GAAP) endow managers with some discretion in their choice of accounting alternatives. The same private information regarding future outcomes can be conveyed differently by different managers, or by the same manager under differing incentives. In such an environment, individuals interpret accounting announcements heterogeneously and gather information to help them assess the implications of accounting announcements. They understand that the announcements may diverge from the manager's private information. While these divergences may result from the strategic decisions implicit in all disclosures, individuals cannot typically observe the specific factors that influence such strategic decisions in order to infer the directions and magnitude of the "bias." Therefore, such strategic disclosure considerations create a "distortion noise" that hampers the interpretation of private information conveyed by managers in accounting announcements; the "distortion noise" is a perception of likely differences between the disclosed information and the actual information observed privately by managers. In sum, users are likely to infer different signals from the disclosure if they do not have the same resources and information to evaluate it.

The noisy rational expectation equilibrium developed in this paper recognizes this characteristic of the accounting information environment by incorporating divergent interpretations of accounting announcements. Doing so allows us to define the *IC* in a way that is consistent with heterogeneous beliefs and with heterogeneous interpretations of accounting news.

Specifically, we require the *IC* definition to satisfy three criteria:

1. It must reflect a weighted average belief change across all market traders.
2. It must reflect the differing degrees of conviction with which such belief changes are associated. Such degrees of conviction are seen as related to the distortion noise that inheres in the earning announcement.
3. It must be derived from an equilibrium price and trading model, so that it could be stated in terms of empirically observable manifestations; only equilibrium behavior is manifested in observed measures or phenomena.

As suggested in Beaver (1968), market reaction to public information can manifest itself either in a price reaction, a trading volume reaction, or both. Beaver conjectured that price reactions reflect changes in consensus beliefs, whereas trading volume reflects disagreement or divergent interpretations among traders about the implications of the announcement for future payoffs (recent empirical evidence in Bamber and Cheon [1993] supports this conjecture). Trading volume and price reactions to information releases have been investigated by Demski and Feltham (1994), Grundy and McNichols (1989), Holthausen and Verrecchia (1990), Kim and Verrecchia (1991a, 1991b), Pfleiderer (1984), and Varian (1985), among others.[4]

---

[3] Specifically, the FASB states "Accounting information may not represent faithfully what it purports to represent because it has one or both of two kinds of bias. The measurement may be biased, so that the resulting measurement fails to represent what it purports to represent. Alternatively, or additionally, the measurer, through lack of skill or lack of integrity, or both, may misapply the measurement method chosen. In other words, there may be bias, not necessarily intended, on the part of the measurer. Verification implies consensus. Verifiability can be measured by looking at the dispersion of a number of independent measurements of some particular phenomenon. The more closely the measurements are likely to be clustered together, the greater the verifiability of the number used as a measure of the phenomenon" (FASB 1980, pars. 78, 84).

Grundy and McNichols (1989) study effects of the informativeness of the price sequence and trading, while Kim and Verrecchia (1991a) investigate trading that results from differences in the precision of traders' prior information. In contrast, the focus of this paper is how anticipated accounting announcements change individuals' beliefs and informedness in ways that can be related to a meaningful notion of the *IC* of such announcements. Our "equilibrium" definition of *IC* is related to trading volume and price changes, among other factors. Unlike Grundy and McNichols (1989) and Kim and Verrecchia (1991a), we characterize public information in a way that leads to differential interpretations by traders. Before the public announcement, individuals are afforded the opportunity to acquire a second private signal. This approach contrasts with those of Kim and Verrecchia (1991a) and Grundy and McNichols (1989) who investigate public disclosures that are free from distortion noise and are interpreted homogeneously. Partially because of this difference, the information variables that induce trading in the second period in our model are structurally different from those in Grundy and McNichols (1989) and Kim and Verrecchia (1991a).

In Grundy and McNichols (1989), trading is induced by essentially non-concordant beliefs about either the equilibrium price signal or the public signal.[5] They show that when the correlation between the private signal common error and the public signal common error is zero, beliefs are essentially concordant, and no trade occurs in the second period unless the second-period equilibrium price is itself informative. Trading in our model, on the other hand, is not induced by essentially non-concordant beliefs about the public signal, because the private information acquired by traders in the first round of trading contains only an idiosyncratic error (and no common error), which is uncorrelated with the common error contained in the second-period public signal. Rather, it is the divergent interpretations that induce trading in the second period in our model. Traders' beliefs about the public signal are always concordant when they trade in the first period. Were it not for the interpretation signal acquired in the second period, beliefs would be essentially homogeneous in the Hakansson et al. (1982) sense, and *no trading* would occur.

In Kim and Verrecchia (1991a), trading in the second period is induced by the arrival of a public signal, interpreted homogeneously, that narrows the dispersion in traders' preannouncement precision weighted expectations. No trading would occur if there were no differences in the precision of traders' prior private signals. This model contrasts with the setting in our paper which assumes away differences in the precision of traders' private information, and instead considers the trading volume reaction to be attributable to the divergence in traders' interpretations of the disclosure. Under the information structure in our model, it is possible that individuals' posterior expectations could become even *more divergent* after observing the public disclosure.

---

[4] The literature on trading volume is more extensive than the above citations imply. We have limited our discussion of the literature to studies that employ the rational expectations equilibrium framework.

[5] Concordant beliefs in the Milgrom and Stokey (1982) sense means that traders do not expect the anticipated signal to change the existing configuration of beliefs. In Grundy and McNichols (1989), a sufficient condition for concordance (for an anticipated public signal) is when common errors in traders' predisclosure private signals are uncorrelated with the error in the public signal.

In a single-trading-round setting, Holthausen and Verrecchia (1990) explored the relation between informedness (the inverse of the variance of the random outcome conditioned on private information and price) and consensus (the correlatedness among the differences between the random outcomes and the conditioned [on announcement] expectations thereof, measured *ex ante*) on one hand, and the variance of the price changes and the expected trading volume on the other. While the private information acquisition model in Holthausen and Verrecchia (1990) may be similar to our model of public information, the other elements of our model differ sufficiently from theirs to make the respective discussions and conclusions incomparable.

The character of the IC of *anticipated* public announcements can be examined by investigating how such disclosures alter, in the aggregate, existing equilibrium beliefs, and by assessing the extent to which the changes in these beliefs are reflected in empirically observable phenomena such as trading and price changes. Holthausen and Verrecchia (1990) discuss separately the impact of informedness and consensus on expected trading volume and the variance of price changes. Because they analyze a single period in which traders start with homogeneous expectations that begin to diverge upon the acquisition of private signals, their analysis does not examine belief changes from an equilibrium of heterogeneous beliefs formed in a preceding round of trading that anticipates the public signal. Unlike Holthausen and Verrecchia (1990), we have a public signal that updates *prior heterogeneous* beliefs just as accounting releases actually do. Such a model setting is essential for the formulation of an operational definition of IC. Our paper considers two trading periods and three events of information acquisition: the first is a private information acquisition in the first period, the second is the acquisition of another private signal for interpreting an anticipated public signal known to contain distortion noise, and the third is the heterogeneous inferences from the announcement. This scheme requires a multiperiod model where trading positions and other information gathering activities undertaken by traders can be taken into account. This model is especially important in our setting where traders do not expect to interpret the public announcement homogeneously. The IC formulation that is derived from our equilibrium model incorporates explicitly prior dispersion in beliefs.

Our model incorporates a characterization of differential inferences (dispersion in interpretations) from public announcements by traders based on *idiosyncratic errors* in inferences from public announcements which typically arise from the accounting environment described above. This explanation differs from Holthausen and Verrecchia's (1990) definition of consensus based on *residual errors* in traders' expectations (the difference between their conditional expectations and the realization of the random outcome). The two measures describe different attributes of the traders' information environment and are not directly comparable. For example, while the residual errors defined in Holthausen and Verrecchia (1990) are correlated due to common errors in individuals' private signals, the idiosyncratic errors in our definition are uncorrelated. Moreover, our measure of interpretation heterogeneity is derived from aggregating differences in traders' inferences from the public disclosure while Holthausen and Verrecchia's (1990) definition of consensus beliefs is based on the correlation in the residual errors. The two definitions are therefore structurally different, and the relation between them and the attributes of the equilibrium are different as well. For example, while Holthausen and Verrecchia (1990) find expected trading volume to be decreasing monotonically in consensus, we find a non-monotonic relation where trading increases initially to a maximum and decreases thereafter as dispersion increases.

II. Analysis and Results

We use a two-period, three-date, noisy rational expectations equilibrium model with a continuum of infinite market traders indexed i and distributed on the unit interval (0, 1). The model is similar in structure to that discussed in Kim and Verrecchia (1991a) and Grundy and McNichols (1989) but it includes heterogeneous interpretations of the public announcement. Each individual maximizes the expected utility of his final period (time 3) wealth, W_{i_3}, that consists of the initial endowment, B_i, plus the returns on his optimal holdings of the risky asset during periods 1 and 2, z_{i_1} and z_{i_2}, expressed as

$$W_{i_3}=B_i+z_{i_1}(P_2-P_1)+z_{i_2}(x-P_2).$$

Starting with a riskless asset endowment B_i, each individual acquires private information about the liquidating value of the risky asset x in the first period and trades anticipating the future arrival of a public disclosure that will alter his beliefs about x resulting in another round in trading. Unlike those in existing models, traders in our model do not interpret the public announcement (η) homogeneously because they are differentially informed about the distortion noise (ν) contained in it. Aware of such noise, traders acquire additional private information, h_i, to help them interpret the disclosure when released. Differences in the acquired signal h_i produce different inferences (y_{i_2}) from the disclosure. The ability to filter out the distortion noise and make inferences about the true signal y depends upon the precision of h_i.[6]

While inferring y_{i_2} from the public disclosure is conceptually similar to the acquisition of an ordinary private signal at time 2, it is quite different because the time at which η will be disclosed is anticipated by all traders, thereby affording the opportunity to take trading positions and gather additional information in preparation for its arrival. Also, since the inferences y_{i_2} are generated from the same disclosure, they will be correlated through the error contained in the public signal. The correlation among and precision of the inferences y_{i_2} depend on the distortion noise ν associated with the disclosure. Thus, y_{i_2} are sufficiently distinguishable from ordinary private information acquisition to warrant an investigation of their attributes and of their relation to the ability to assess the IC of accounting announcements.

---

[6] The public disclosure is $\eta(y)=y+\nu$, where $y=x+u$ (u being a common random noise) is the imperfect underlying signal about x, and ν is the distortion noise. The coefficients x, u, and v are normally distributed and mutually independent, with zero means and variances σ_x^2, σ_u^2, and σ_ν^2, respectively (assuming zero mean in the case of x does not involve loss in generality). The coefficient h_i reflects the true distortion component inherent in the disclosure with an additive random noise n_i also distributed normally with zero mean and a finite variance $\sigma_{n_i}^2$, i.e., $h_i=\nu+n_i$. We assume that h_i provides information only about the distortion noise associated with the disclosure and does not provide additional information about the outcome. This assumption is needed to rule out another round of trading in anticipation of the disclosure. Since h_i is privately acquired, the filtering out of the distortion noise and the inferences from the disclosure will differ among individuals. We model the filtration process as $E(\nu|h_i)=\delta h_i$, where $\delta=\sigma_\nu^2/(\sigma_\nu^2+\sigma_n^2)$ which is subtracted from the disclosure to yield the inferred signal $y_{i_2}=y+[\nu-E(\nu|h_i)]=y+\xi_i$, where $\xi_i=\nu-\delta h_i=(1-\delta)\nu+\delta n_i$ and δn_i is the residual error in the individual's inference from the disclosure.

Upon inferring y_{i_2}, traders revise their expectations about the mean and variance of the outcome of the risky asset x, and rebalance their first period portfolios through a second round of trading. Solving the optimal holding of the risky asset by backward optimization, each trader uses four potential information sources: the private information acquired in the first period y_{i_1}; inferences from the public announcement in the second period y_{i_2}; the observed first period equilibrium price P_1; and the observed prevailing second period equilibrium price P_2. The price function is conjectured to be linear in the underlying random variables x, y, Z_1, and Z_2, where Z_1 and Z_2 (both i.i.d. with mean zero and variance σ_z^2) are the per capita supply of risky assets in period 1 and period 2, respectively. This four-tuple is denoted $\Phi_{i_2}=\{y_{i_1},y_{i_2},P_1,P_2\}$. We can show that:

$$z_{i_2}=\frac{1}{V_{i_2}}[E(x|\Phi_{i_2})-P_2]=Q_{21}(y_{i_1}-\bar{y}_1)+Q_{22}(y_{i_2}-\bar{y}_2)+Z_2, \tag{1}$$

where:

$$V_{i_2}=\mathrm{var}(x|\Phi_{i2})$$

$$Q_{21}=\frac{B_{21}}{V_2},\ Q_{22}=\frac{B_{22}}{V_2},\ Q_{23}=\frac{B_{23}}{V_2},\ Q_{24}=\frac{B_{24}}{V_2},\ \text{and}\ Q_{25}=\frac{1}{V_2}-Q_{24},$$

and where B_{21}, B_{22}, B_{23}, and B_{24} are the coefficients of the conditioning information variables in the expectation:

$$E(x|\Phi_{i_2})=B_{21}y_{i_1}+B_{22}y_{i_2}+B_{23}P_1+B_{24}P_2,$$

where:

$$\bar{y}_2=\int_0^1 y_{i_2}di$$

is the average inference from the disclosure. V_{i_2} is the variance of x conditional on trader i's time 2 information set Φ_{i_2}. Information precision and risk tolerances are assumed to be identical across individuals and therefore $V_{i_2}=V_2$ for all i.[7]

Given this solution for the optimal demand for period 2, we solve for the optimal demands in time 1 to obtain:

$$\begin{aligned} z_{i_1}&=\frac{1}{k}\{[E(P_2|y_{i_1},P_1)-P_1]+t[E(x|y_{i_1},P_1)-E(P_2|y_{i_1},P_1)]+l(y_{i_1}-\bar{y}_1)\}\\ &=r(y_{i_1}-\bar{y}_1)+Z_1, \end{aligned} \tag{2}$$

where:

$$\bar{y}_1=\int_0^1 y_{i_1}di$$

---

[7] The derivations of the demand functions involving mostly technical algebraic manipulations have been omitted for space considerations. They can be furnished by the authors upon request.

is the average first period signal; the coefficients k, t, and l are closed form functions of variances of the information variables y_{i_1} and y_{i_2}, and the random per capita supply Z_1; and

$$r = aB_{11} + c,$$

where B_{11} is the coefficient of the private signal y_{i_1} in the conditional expectation:

$$E(x \mid y_{i_1}, P_1) = B_{11} y_{i_1} + B_{12} P_1;$$

$$a = \frac{1}{k}\left[t + (1-t)\left(\frac{Q_{21}}{Q_{25}} + \frac{Q_{22}}{Q_{25}}\right)\right], \text{ and}$$

$$c = \frac{l}{k}.$$

Traders' asset demands in each period thus depend simply on the idiosyncratic component in private signals (the difference between the private and the average signal), and on the per capita asset supply.

We develop an operational definition of public disclosure *IC* that is consistent with the above equilibrium demand functions and that satisfies the criteria specified above. Additionally, it is consistent with the attributes of financial information considered desirable by rule making bodies such as the FASB. The definition is consistent with an equilibrium pricing mechanism, so that it can, *in principle*, become observable in empirical regularities that can be used to infer the underlying *IC* at least in an ordinal sense. The definition that emerges from our equilibrium analysis demonstrates that the empirical construct usually employed must be supplemented to gauge the "real content" or impact of the announcement. A suitable definition should reflect not only the changes in individuals' beliefs, but also the degree of conviction with which such beliefs are held or changed as a result of the announcement. To reflect these attributes, we define individual *i*'s beliefs at time *t* (β_{t_i}) as a function of the expectation of the liquidating value of the risky asset x, conditional on his information available at the time, Φ_{i_t}; or, $\beta_{t_i} = E(x \mid \Phi_{i_t})$. An individual's degree of conviction, Θ_{i_t}, (how well informed he is) is determined by the dispersion of his subjective probability distribution over possible realizations of x and other random variables. Consequently, we define the individual's degree of conviction in period 2 as $\Theta_{i_2} = 1/\text{var}(x \mid \Phi_{i_t})$, and $\Theta_{i_1} = 1/a$, where a, defined in equation (2) above, reflects the perceived uncertainty of the first period associated with price P_2 of the second period as well as x. The definition also accommodates an environment with multiple decision makers, each of whom makes a different inference from the announcement. Incorporating all these factors yields a definition of *IC* that reflects aggregate absolute changes in individuals' beliefs weighted by their associated degrees of conviction expressed as:

$$IC \equiv \int_0^1 \left| \Theta_{i_2}\beta_{2_i} - \Theta_{i_1}\beta_{1_i} \right| di.$$

Expected *IC* defined as aggregate expected absolute changes in beliefs and their associated degrees of conviction, takes the form,

$$E(IC) = \int_0^1 E(\left| \Theta_{i_2}\beta_{2_i} - \Theta_{i_1}\beta_{1_i} \right|) di.$$

Assuming no supply shocks (i.e., changes in the supply of the risky asset attributable to liquidity trading), or alternatively controlling for them, expected *IC* is, in equilibrium:

$$E(IC)=(2\pi)^{1/2}\left\{[f(P_2-P_1)]^2+2[E(VOL)]^2+\frac{1}{2}\alpha\int_0^1 E(y_{i_1}-\bar{y}_1)^2 di\right\}^{1/2} \tag{3}$$

where:

$$E(VOL)=\frac{1}{2}\int_0^1 |z_{i_2}-z_{i_1}|di$$

is the expected trading volume associated with the disclosure; z_{i_1} and z_{i_2} are *i*'s equilibrium demands of the risky assets at time 1 and 2;

$$f(P_2-P_1)=\frac{1}{2}E(|\Theta_2 P_2-\Theta_1 P_1|)$$

is the weighted change in price, monotone increasing in the actual price change P_2-P_1;

$$\int_0^1 E(y_{i_1}-\bar{y}_1)^2 di=\mathrm{var}(e_i)$$

is the expected dispersion in individuals private signals held prior to the public disclosure; and

$$\alpha=2c(Q_{21}-r)+c^2$$

depends upon the variance-covariance matrix of the random variables (see eq. [2]).[8]

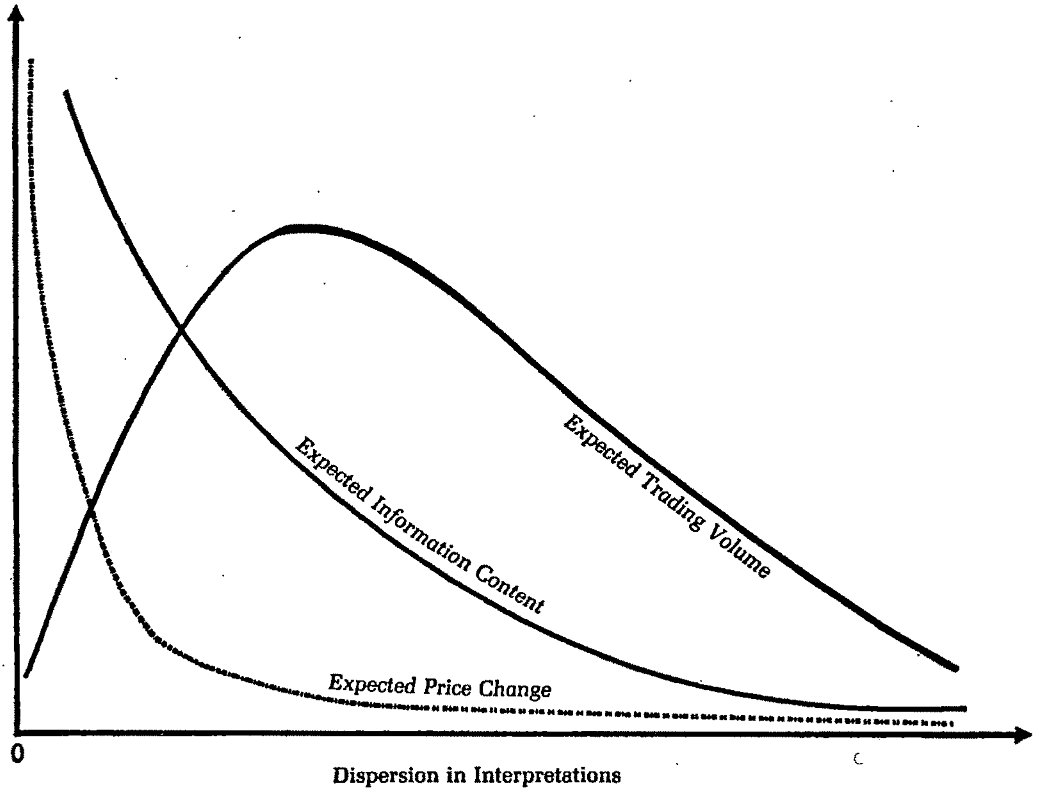
III. Discussion and Conclusions

Our results show that the *IC* of an accounting announcement is directly related to the trading volume, the price reaction associated with the announcement, and the expected divergence in prior private signals. Thus, the information environment existing prior to the announcement must also be explicitly considered to gauge *IC*. Specifically, as is evident from equation (3), the *IC* of the disclosure takes into account the dispersion in traders' beliefs prior to its arrival (which could be proxied by prior analysts' forecast dispersion). The necessity to consider prior dispersion in beliefs demonstrates that our "content" definition has profound implications that go beyond the necessity of *observing jointly* price changes and trading activity. *IC* is mostly higher with higher prior dispersion for a given combination of trading activity and price change. The intuition is that to effect a given magnitude of price change and trading activity, beliefs must change by a well-defined amount. Whether beliefs converge or diverge will not significantly affect the combined magnitudes of price change and trading activity provided the *configuration* of belief changes is preserved. For example, in the stylized case of two traders, the combined variables of trading and price changes may be affected equivalently by a change from traders' divergent first period beliefs of 8 and 12 to identical

---

[8] Proof of the relation is omitted for space considerations. It is available from the authors upon request.

Figure 1
Effect of Dispersion in Public Disclosure Interpretations on Information Content, Trading Volume, and Price Change



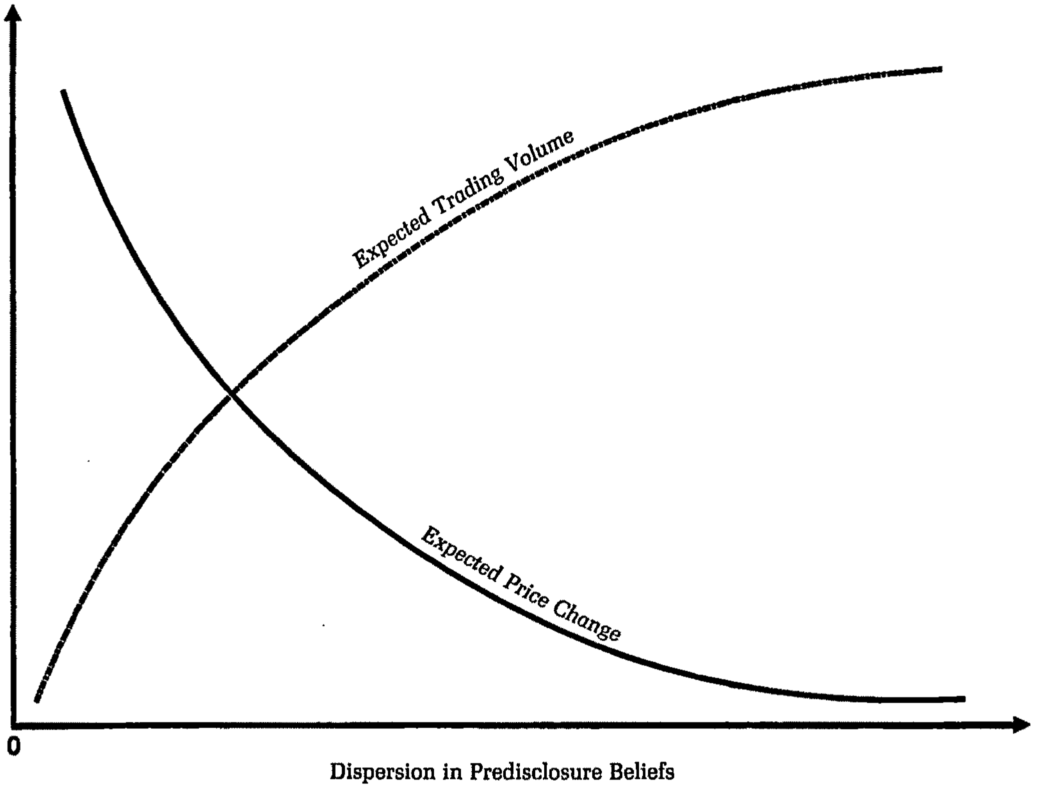
estimates of 11 and 11 (where the numbers reflect the conditional expectations of the outcome of the risky asset in the first and second period, respectively), and by a change in beliefs from identical estimates of 11 and 11 to divergent estimates of 8 and 12. However, a "stronger" public signal—"stronger" in the sense of overall precision (the precision of the manager's private signal and the inverse of the variance of the distortion noise)—would be required to induce the traders to converge from divergent prior beliefs of 8 and 12 to posterior beliefs of 11 and 11 than to make them diverge from identical prior beliefs of 11 and 11 to posterior divergent beliefs of 8 and 12. Hence, the prior precision is a vital measure to consider in properly gauging IC since trading and price changes do not capture the full "content" or "quality" of the disclosure that inheres in its ability to bring beliefs together as opposed to spreading them apart.

Numerical analysis reveals that while expected IC decreases with the dispersion of individuals' inferences from the disclosure, it increases along with the price reaction with the precision of the announcement.[9] It is also interesting to note, as figure 1 demonstrates, that there exists a non-monotone relation between expected trading volume and the dispersion of traders' inferences from the announcement. This conclusion is consistent with empirical evidence in Ajinkya et al. (1991) that shows that trading volume is positively correlated with the dispersion in analysts' forecasts. Expected trading volume increases initially up to a maximum and decreases thereafter as dispersion increases. We attribute this behavior to countervailing effects of disclosure noise on trading volume. On one hand, larger distortions in the disclosure make it more likely that traders will draw different inferences from it, inducing more widely divergent expectations about the firm's future outcomes and increased trading. On the other hand, a large distortion in the disclosure implies more uncertainty about its implications for future outcomes and a lower *IC*. This result reduces the extent of belief revision the disclosure induces among traders, thereby weakening incentives to trade. A maximum point is reached when the heightened uncertainty in traders' expectations offsets the incentives to trade generated by the divergence in their inferences. Beyond this point expected trading declines as the uncertainty about the *IC* of the announcement more than offsets trading induced by interpretation divergence. Figure 2 shows how the information environment that existed prior to the announcement (i.e., the existing quality and dispersion in individuals private information) affects the *IC* of public announcements. It shows that trading volume increases and price change decreases with expected dispersion in predisclosure beliefs. This is generally consistent with empirical evidence presented by Ajinkya et al. (1991) and Bamber and Cheon (1993), who find that price reaction is lower and trading volume higher when beliefs are more generally dispersed.

These results are important because they demonstrate the need *jointly* to observe *both* the price reaction *and* the trading activity coincident with the announcement (the price change may underestimate *IC* when observed in isolation) and the information environment that existed prior to the announcement in order to assess the announcement's content.

[9] In conducting the numerical analysis, all parameters and variables including the aggregate supply parameters, Z_2 and Z_1, were set to unity except for the parameters under investigation.

Figure 2
Effect of Dispersion in Predisclosure Beliefs on Price Reaction and Expected Trading Volume



References

Ajinkya, B., R. Atiase, and M. Gift. 1991. Volume of trading and the dispersion of financial analysts' earnings forecasts. *The Accounting Review* 66 (April): 389–401.

Bamber, L., and S. Cheon. 1993. Differential price and volume reactions to annual accounting earnings announcements. Working paper, University of Georgia.

Beaver, W. 1968. The IC of annual earnings announcements. *Journal of Accounting Research* (Supplement): 67–92.

Bernard, V. L. 1992. Stock price reactions to earnings announcements: A summary of recent anomalous evidence and possible explanations. In *Advances in Behavioral Finance* edited by R. Thaler. New York: Sage Publications.

Demski, J., and G. Feltham. 1994. Market response to financial reports. *Journal of Accounting & Economics* 17. Forthcoming.

Fama, E. F. 1991. Efficient capital markets: II. *Journal of Finance* 46 (December): 1575–1617.

Financial Accounting Standards Board. 1980. Statements of Financial Accounting Concepts No. 2: *Qualitative Characteristics of Accounting Information*. Stamford, CT: FASB.

Grossman, S., and J. E. Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70 (June): 393–408.

Grundy, B., and M. McNichols. 1989. Trade and the revelation of information through prices and direct disclosure. *Review of Financial Studies* 2: 495–526.

Hakansson, N., G. Kunkel, and J. Ohlson. 1982. Sufficient and necessary conditions for information to have social value in pure exchange. Journal of Finance (December): 1169–81.

Hirshleifer, J. 1973. Where are we in the theory of information? *American Economic Review* Proceedings 63 (May): 31–39.

Holthausen, R., and R. Verrecchia. 1990. The effect of informedness and consensus on price and volume behavior. *The Accounting Review* 65 (January): 191–208.

Karpoff, J. M. 1987. The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis* 22 (March): 109–26.

Kim, O., and R. Verrecchia. 1991a. Trading volume and price reactions to public announcements. *Journal of Accounting Research* 26 (Autumn): 302–21.

——, and ——. 1991b. Market reaction to anticipated announcements. *Journal of Financial Economics* 30 (December): 273–309.

Milgrom, P., and N. Stokey. 1982. Information, trade and common knowledge. *Journal of Economic Theory* 26 (February): 17–27.

Pfleiderer, P. 1984. The volume of trade and the variability of prices: A framework for analysis in noisy national expectations equilibria. Working paper, Stanford University.

Theil, H. 1967. Economics and information theory. Stokie, IL: Rand McNally.

Varian, H. R. 1985. Differences on opinion and the volume of trade. Working paper, University of Michigan.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 870–873

How Do We Assess a Model of Price and Volume?

Robert E. Verrecchia
University of Pennsylvania

IN this comment I discuss the paper "Information Content of Accounting Announcements" by Alex Dontoh and Joshua Ronen.[1] While the stated purpose of Dontoh and Ronen (1993) is to offer an operational definition of "information content" in the context of a competitive, rational expectations model of sequential trade, the recent proliferation of similar-type models requires some discussion of purely modeling issues to start. The feature common to all the competitive, rational expectations models is that they exploit a technology that preserves linearity: negative exponential utility functions, normally distributed random variables, linear conjectures about the behavior of prices, etc. To a naive reader, this practice may create the impression that all these models are equivalent. Despite the use of common technology, this impression is incorrect because the characterizations of price and volume that result from each model are highly sensitive to the information structure assumed to underlie the economy. Examples of different information structures include: private information with or without common error, private information with identical precision of idiosyncratic error terms versus different precisions, and public announcements that are interpreted either homogeneously or heterogeneously. In effect, each information structure (potentially) implies a unique price and volume characterization. In addition, the sequence in which traders receive private information in relation to public announcements also potentially affects the underlying information structure, thus adding to the possibilities. These variations imply that, at least in theory, there exists a multiplicity of ways of creating different information structures, yielding a veritable cornucopia of (potentially different) price and volume characterizations. What is a reader to make of this?

Implicitly, Dontoh and Ronen (1993) attempt to distinguish their contribution through three arguments: first, their information structure is different from previous

---

[1] I was invited to write a comment on "Information Content of Accounting Announcements" by the current editor of *The Accounting Review*, A. Rashad Abdel-khalik. However, I was not involved in the decision to publish "Information Content of Accounting Announcements" in the journal, and thus this comment should be interpreted neither as an endorsement or a criticism of that decision. Also, I would like to acknowledge editorial suggestions from Ann Beatty in preparing this comment.

Invited by the Editor.
Received May 1993.

ones, which makes their contribution unique; second, their information structure is more appropriate than ones previously suggested; and third, their information structure leads naturally to a useful definition of information content.

With regard to the first argument, while it is true that Dontoh and Ronen's (1993) choice of an information structure is unique, this claim may miss the point that the role of these models should not be to document the price and volume relation implied by every conceivable information structure, but rather to offer compelling explanations of the economics underlying these relations. For example, previous literature has offered the following economic intuition underlying price and volume relations. Price is a type of aggregator of information that averages the opinions and information of investors in the marketplace, giving weight to each opinion on the basis of the confidence expressed by that investor through the aggressiveness of his or her demand for the asset. Through this averaging process, price yields a type of "consensus judgment" that ameliorates the problem created by the idiosyncratic error in each individual investor's opinion. Volume, on the other hand, arises from differences of opinion. While volume is driven in part by the same shift in opinion captured by a change in price, volume aggregates the idiosyncratic error in individual investors' opinions (as opposed to averaging it out, as in the case of price). Idiosyncratic errors exacerbate the volatility of volume more so than price. Therefore, while price (or returns) can be thought of as a measure of the market's consensus judgment about the information content of an accounting announcement, volume can be thought of as a noisier measure of the same phenomenon, noisier because of the greater role played in volume's determination by the idiosyncratic errors in investors' opinions. That price change and volume capture some common aspect of the information content of an accounting announcement is manifest in theoretical characterizations by the fact that price changes and volume are positively related, a point also well-established in the empirical literature. Rather than simply consider alternative, and more complex, information structures, a useful contribution to the literature should extend our understanding of the economic intuition underlying price and volume relations beyond those ideas that are outlined above.

With regard to the second argument, the ultimate arbiter of "appropriateness" is confirmation through empirical investigation. While more complicated models of trade that extend our economic intuition will always be welcomed, important advances are more likely to come from linking existing models of rational trade to testable, empirical observations. Through empirical investigation, we can confirm the appropriateness of the underlying assumptions directly. For example, if volume is a noisier measure of the impact of an accounting announcement than price, and price and volume are the only economic statistics that are directly observable, what do we learn from volume in combination with price change that we do not learn from price alone? Appealing to the intuition offered in Kim and Verrecchia (1991a, 1991b), Stickel and Verrecchia (1993) argue that, holding price change constant, volume measures the extent of private information gathering in anticipation of a public announcement. Since bid-ask spreads proxy for the amount of private information in the market, holding price change constant, unexpected trading volume is positively related to the subsequent change in the bid-ask spread. In addition, holding price change constant, unexpected trading volume, as a proxy for private information, distinguishes information-motivated trade from liquidity-motivated trade, and thus permanent price changes from temporary price changes. (These hypotheses are then tested using quarterly earnings announcements, and evidence consistent with these predictions is found.) This work is very preliminary, however, and the role volume plays in conjunction with price change toward understanding the impact of an accounting announcement is an issue that requires further attention.

Furthermore, in light of very recent work, the information structure employed in Dontoh and Ronen (1993) has an obsolescent feature, which limits the paper's application. In Dontoh and Ronen (1993) the acquisition of private information in anticipation of a public announcement, as well as the information to interpret the public announcement itself, is treated as exogenous. When public announcements are anticipated, the central issue is how their anticipation affects the decision to acquire private information in advance of the announcement. This decision, in turn, affects the behavior of price changes and volume subsequent to the announcement being made. Modeling explicitly the decision investors make to acquire private information may substantially change price and volume behavior. For example, an interested reader could compare the results in Kim and Verrecchia (1991a) with Kim and Verrecchia (1991b), and see that the shift from exogenous information acquisition in the former paper to endogenous information acquisition in the latter can change the behavior of price and volume as a function of other parameters (e.g., the precision of the public announcement). In the context of a competitive, rational expectations model of sequential trade, endogenous information acquisition has been discussed extensively in Kim and Verrecchia (1991b) and Demski and Feltham (1994). Thus, the continued usefulness of models based on exogenously specified information structures is limited.

A discussion of the usefulness of the Dontoh and Ronen (1993) analysis inevitably leads to a consideration of the primary focus of Dontoh and Ronen (1993), namely, providing an operational definition of information content. First, note that all suggested measures of information content are inherently arbitrary, and this characteristic calls into question whether any single one can be promoted as "better" than some other. Furthermore, the criteria outlined in Dontoh and Ronen (1993) for a definition of information content are sufficiently broad so as to include price change and volume as measures of information content in all previous models of rational trade. While there may exist a variety of alternative possible candidates to measure information content, only price change and volume require no proxies, and in this sense are truly "operational." For example, Dontoh and Ronen's (1993) definition of information content depends on the expected dispersion of traders' private signals held prior to a public announcement, which is unobservable. (There are other parameters in Dontoh and Ronen's definition for which proxies might need to be obtained, but brevity precludes a detailed discussion of each.) The disadvantage of proxies in a definition of information content, versus using price change and volume directly, is that they add noise. Furthermore, even in the event that suitable proxies can be found for some of these parameters, the definition itself remains sensitive to the economy's assumed underlying information structure, for which no proxy would seem to exist. Consequently, as a practical matter, it is difficult to imagine that one could improve on the use of price change alone or price change in combination with volume as operational measures of information content.

References

Demski, J. S., and G. A. Feltham. 1994. Market response to financial reports. *Journal of Accounting & Economics* 17. Forthcoming.

Dontoh, A., and J. Ronen. 1993. Information content of accounting announcements. *The Accounting Review* 68 (October): 857–69.

Kim, O., and R. E. Verrecchia. 1991a. Trading volume and price reactions to public announcements. *Journal of Accounting Research* 26 (Autumn): 302–21.

——, and ——. 1991b. Market reaction to anticipated announcements. *Journal of Financial Economics* 30 (December): 273–309.

Stickel, S. E., and R. E. Verrecchia. 1993. Evidence that volume sustains price changes. Working paper, University of Pennsylvania.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 874–884

Information Acquisition in a Tax Compliance Game

Richard C. Sansing
Yale University

SYNOPSIS AND INTRODUCTION: The Internal Revenue Service (IRS) relies increasingly on its ability to detect taxpayer noncompliance without engaging in a comprehensive individual audit. The IRS's compliance initiative, Compliance 2000, emphasizes the targeting of noncompliant taxpayers rather than relying on random audits to enforce the tax laws. For example, the IRS uses a model developed from the Taxpayer Compliance Measurement Program (TCMP) to help it choose which returns to audit.

The treatment of losses from tax shelter partnerships presents a difficult compliance problem for the IRS. It is not evident from the face of either the partnership return or the partner's return whether the losses from the partnership can be legitimately deducted. A plausible audit strategy is for the IRS to develop models that can predict when an individual is improperly deducting a loss. The tax shelter disclosure rules in I.R.C. §6111 and §6112 provide information to the IRS that helps it detect taxpayers investing in abusive tax shelters.

Previous work has modeled tax compliance as a game between a wealth-maximizing taxpayer and a tax enforcement agency trying to maximize government revenues, net of audit costs (Graetz et al. 1986; Reinganum and Wilde 1986; Beck and Jung 1989). In these papers, the IRS uses the taxpayer's declaration of income when it decides whether to audit that taxpayer.

The purpose of this paper is to examine the effect of information that helps the IRS predict tax evasion on the strategic choices made by the taxpayer and the IRS. The information has a *direct* effect by giving the IRS information that can improve its audit decision. It also has an *indirect* effect by changing the taxpayer's incentives to engage in tax evasion, which in turn changes the IRS's incentives to audit taxpayers. The optimal level of information acquisition is also examined.

The analysis yields four important results regarding the effect of information on tax compliance. First, it can induce an increase in tax evasion.

I thank Amir Ziv, Joel Demski, workshop participants at New York University and the University of Michigan, and three anonymous reviewers for helpful comments.

Submitted September 1992.
Accepted April 1993.

Second, it has no effect on the expected level of gross government revenues. Third, it can increase expected audit costs. Fourth, the optimal level of investment in information acquisition does not vary monotonically with tax rates, penalty rates, audit costs, or the amount of loss deducted by the taxpayer.

Key Words: *Tax compliance, Information economics, Strategic auditing.*

SECTION I presents the model. In section II, I characterize the equilibrium and derive comparative statics. In section III, I examine the value of the signal to the IRS and the optimal level of information acquisition. Section IV contains conclusions.

I. Model

There are two players, a taxpayer (T) and the IRS. The taxpayer observes his tax status (u,v,λ). The variable u represents the taxpayer's income from salary and investments other than tax shelters. The variable v represents the taxpayer's distributive share of losses from a tax shelter partnership. The joint distribution of u and v is arbitrary. The variable $\lambda \in (0,1)$ indicates whether the taxpayer can legally deduct v on his tax return. If $\lambda = 1$, the loss is not deductible either because the investment lacks economic substance or because partner-level restrictions such as the at-risk limitations under §461 or the passive activity limitations under §469 disallow the deduction; if $\lambda = 0$, v is currently deductible. The amounts u and v are reported to the IRS via W-2, 1099, and K-1 forms. In contrast, λ is observed privately by the taxpayer. Thus, the IRS observes the amount of the taxpayer's loss but does not observe the correct tax treatment of the loss. Both the taxpayer and the IRS know $\text{Prob}(\lambda = 1 | u,v) = q$.

The taxpayer files a tax return $\theta \in (u-v,u)$. For expositional convenience, taxpayers for whom v is deductible are referred to as low types; others are referred to as high types. Similarly, a taxpayer who deducts v is said to file a low return, and others are said to file a high return.

The IRS costlessly observes u,v,θ, and a signal z, which is correlated with λ. The signal is the output from the IRS's tax evasion prediction model. This model is developed with data from tax shelter registration disclosures, prior audit experience, and TCMP audits, which are comprehensive audits of randomly selected tax returns. The signal is modeled as a realization from a logistic distribution with mean $\lambda\eta$, where the parameter $\eta \geq 0$ is a measure of the signal's precision, that is, the extent to which it is correlated with the correct tax treatment. After observing the signal, the IRS chooses an action, $\omega \in (Y,N)$, where Y indicates that the IRS audits the taxpayer and N indicates that the IRS does not audit.

The IRS is of course very secretive regarding the correlation between the information it receives and the frequency of tax evasion. On the other hand, aggregate level data such as the percentage of returns audited and the average amount of taxes and penalties collected from audits are reported publicly in *The Annual Report of the Commissioner of Internal Revenue*. In addition, tax practitioners are a source of information about audit probabilities. Thus I assume that taxpayers know η, the average ability of the IRS to predict a taxpayer's type, but do not know z, the prediction in their particular case.

The taxpayer and the IRS are risk-neutral with von Neumann-Morgenstern utility functions $U^T(u,v,\lambda,\theta,\omega)$ and $U^{IRS}(u,v,\lambda,\theta,\omega)$. The taxpayer owes a tax of $t(u-v+\lambda v)$, where $t>0$ is the tax rate. When a taxpayer files $\theta=u$, he pays a tax of tu; when the taxpayer files $\theta=u-v$ and is not audited, he pays $t(u-v)$. If the IRS conducts an audit, the taxpayer pays the correct tax due plus a penalty proportional to the understated tax.[1] The penalty rate is p, so if a high type files a low return and is audited, the total payment to the IRS is $t(u+pv)$.[2] If an audit occurs, the IRS incurs a cost $c>0$. The cost of auditing must be low enough so that the threat of the IRS to audit is sufficiently credible to deter some underreporting. Thus, I make the following assumption:

$$c<\frac{qtv(1+p)(1+pe^{\eta})^2}{(1-q)e^{\eta}(1+p)^2+q(1+pe^{\eta})^2}.$$

The intuition behind the upper bound on c becomes clearer by examining the case $\eta=0$. Then the condition becomes $c<qtv(1+p)$, which implies the IRS prefers to audit a return if all taxpayers deduct their loss, even if the IRS has no ability to predict whether a loss is properly deductible. Note the upper bound on c implies $tv(1+p)>c$; if the IRS knows tax evasion has occurred, it prefers to audit. There are eight possible outcomes, the payoffs from which are shown in table 1.

A solution to this game is a Nash equilibrium. A Nash equilibrium is defined as a strategy for each player where each player's strategy is a best response to the other player's strategy. Formally, a Nash equilibrium in this game is defined as a pair of probability distributions over θ and ω, $P^*[\theta(u,v,\lambda)]$ and $Q^*[\omega(u,v,\theta,z)]$ such that,

(1) for each $(u,v,\lambda), P^*[\theta(u,v,\lambda)]$ solves,

$$\max_{P[\theta(u,v,\lambda)]} \sum_{\theta}\sum_{\omega}\int_{z} [U^T(u,v,\lambda,\theta,\omega)][\mathrm{Prob}(z|u,v,\theta)] \, Q^*[\omega(u,v,\theta,z)]P[\theta(u,v,\lambda)]dz, \text{ and}$$

(2) for each $(u,v,\theta,z), Q^*[\omega(u,v,\theta,z)]$ solves,

$$\max_{Q[\omega(u,v,\theta,z)]} \sum_{\omega}\sum_{\lambda}[U^{IRS}(u,v,\lambda,\theta,\omega)]\mathrm{Prob}(\lambda|u,v,\theta,P^*,z)Q[\omega(u,v,\theta,z)].$$

II. Equilibrium

The IRS observes z, u, v, and θ prior to making an audit decision. The probability that the IRS makes an incorrect decision depends on the quality of the signal, the taxpayer's reporting strategy, and the IRS's decision rule, $\delta, -\infty<\delta<\infty$. Each type λ induces a logistic distribution with mean $\eta\lambda$.[3] The logistic distribution has a probability density,

$$f(z)=\frac{e^z}{(1+e^z)^2},$$

---

[1] This "perfect investigation" assumption condition is not crucial to the paper's fundamental results. All that is required is that the taxpayer would prefer to report truthfully if he knew the IRS were going to audit.

[2] The assumption that the penalty is proportional to the tax liability is consistent with most penalty provisions in the U.S. Interal Revenue Code. See I.R.C §6651, §6662, and §6663.

[3] The important characteristics of this distribution are that its support is the same for both high and low types and that it exhibits the monotone likelihood ratio property (MLRP). Given any nondegenerate prior belief regarding the taxpayer's type, the posterior belief that the taxpayer is a high type is strictly increasing in z.

Table 1
Matrix of Payoffs to the Taxpayer and the IRS

	IRS Action	
Taxpayer Action	*Audit* ($\omega=Y$)	*No Audit* ($\omega=N$)
High type reports high ($\lambda=1,\theta=u$)	$[-tu,tu-c]$	$[-tu,tu]$
High type reports low ($\lambda=1,\theta=u-v$)	$[-t(u+pv),t(u+pv)-c]$	$[-t(u-v),t(u-v)]$
Low type reports high ($\lambda=0,\theta=u$)	$[-t(u-v),t(u-v)-c]$	$[-tu,tu]$
Low type reports low ($\lambda=0,\theta=u-v$)	$[-t(u-v),t(u-v)-c]$	$[-t(u-v),t(u-v)-c]$

Note: Payoffs are in the order [*taxpayer*, IRS].

where z represents the number of standard deviations from the mean; thus λ and η combine to induce a distribution of signals with density,

$$\frac{e^{z-\lambda\eta}}{(1+e^{z-\lambda\eta})^2}.$$

The parameter η is a measure of how well the signal predicts the correct tax treatment. The decision rule is a cutoff point δ, where if $\theta=u-v, \omega=Y$ if $z\geq\delta$, and $\omega=N$ otherwise. This strategy is optimal because the logistic distribution exhibits MLRP. The values of η and δ determine the audit probabilities as follows:

$$\text{Prob}(\omega=Y|\lambda=1)=\int_{\delta}^{\infty}\frac{e^{z-\eta}}{(1+e^{z-\eta})^2}dz=\frac{1}{1+e^{\delta-\eta}} \text{ and} \tag{1a}$$

$$\text{Prob}(\omega=Y|\lambda=0)=\int_{\delta}^{\infty}\frac{e^{z}}{(1+e^{z})^2}dz=\frac{1}{1+e^{\delta}}. \tag{1b}$$

An increase in δ decreases the probability of a type II error (failing to audit a taxpayer who incorrectly deducts the loss) and increases the probability of a type I error (auditing a taxpayer who correctly deducts the loss). The IRS's choice of δ reflects its willingness to trade off type I and type II errors.

The solution to this game is described in Proposition 1.

Proposition 1: The following strategies constitute a Nash equilibrium.

For a taxpayer whose type is $(u,v,0)$ or $(u,v,1)$:

(1a) if the taxpayer's type is low, $\lambda=0$, T reports low, $\theta=u-v$, and
(1b) if the taxpayer's type is high, $\lambda=1$, T reports low, $\theta=u-v$ with probability ϵ^* and reports high, $\theta=u$, with probability $(1-\epsilon^*)$, where,

$$\epsilon^* = \frac{e^{\eta}(1+p)^2 c(1-q)}{(1+pe^{\eta})^2 q[tv(1+p)-c]}.$$

For the IRS:

(2a) if the taxpayer reports high, $\theta = u$, IRS does not audit, $\omega = N$;
(2b) if the taxpayer reports low, $\theta = u - v$, and the signal is above the cutoff, $z \geq \delta^*$, the IRS chooses to audit, $\omega = Y$, where $\delta^* = \eta + \log p$; and
(2c) if the taxpayer reports low, $\theta = u - v$, and the signal is below the cutoff, $z < \delta^*$, the IRS chooses not to audit, $\omega = N$.

Proof: See the Appendix.

The intuition behind Proposition 1 is the decision rule that would be used if all taxpayers deducted v deters high types from doing so. But, if all taxpayers report truthfully, the IRS should never audit; and if the IRS never audits, all taxpayers should deduct the loss. In equilibrium, the IRS adopts the decision rule $\delta^* = \eta + \log p$. This ensures that the probability of a high type that files a low return being audited is,

$$\frac{1}{1+p},$$

so all high types are indifferent between paying a tax of tu and playing the audit lottery that will result in either paying $t(u-v)$ or $t(u+vp)$. High types in turn randomize to make the IRS indifferent between auditing and not auditing when $z = \delta^*$.[4]

The equilibrium is sensitive to the assumptions regarding the game's information structure. If the taxpayer knew z, for example, there would be a critical value z^* for which all types below z^* would report low and the IRS would not audit, and all high types above z^* would randomize to make the IRS indifferent between auditing and not auditing, and the IRS would randomize to make high types indifferent between reports.

The likelihood of the five possible outcomes of this game can be calculated directly from the best response strategies in Proposition 1.

$$\mathrm{Prob}(\lambda=1, \theta=u-v, \omega=Y) \equiv x_1 = \frac{q\epsilon e^{\eta-\delta}}{1+e^{\eta-\delta}} = \frac{e^{\eta}(1+p)c(1-q)}{(1+pe^{\eta})^2[tv(1+p)-c]}, \tag{2a}$$

$$\mathrm{Prob}(\lambda=1, \theta=u-v, \omega=N) \equiv x_2 = \frac{q\epsilon}{1+e^{\eta-\delta}} = \frac{e^{\eta}p(1+p)c(1-q)}{(1+pe^{\eta})^2[tv(1+p)-c]}, \tag{2b}$$

$$\mathrm{Prob}(\lambda=1, \theta=u, \omega=N) \equiv x_3 = q(1-\epsilon)$$

$$= \frac{q(1+pe^{\eta})^2[tv(1+p)-c]-(1-q)(1+p)^2 ce^{\eta}}{q(1+pe^{\eta})^2[tv(1+p)-c]}, \tag{2c}$$

$$\mathrm{Prob}(\lambda=0, \theta=u-v, \omega=Y) \equiv x_4 = \frac{1-q}{1+e^{\delta}} = \frac{1-q}{1+pe^{\eta}}, \text{ and} \tag{2d}$$

$$\mathrm{Prob}(\lambda=0, \theta=u-v, \omega=N) \equiv x_5 = \frac{(1-q)e^{\delta}}{1+e^{\delta}} = \frac{p(1-q)e^{\eta}}{1+pe^{\eta}}. \tag{2e}$$

---

[4] See Shibano (1990) for the characterization of the equilibrium in a strategic auditing game in a more general setting.

Table 2
Effect of Parameter Changes on Outcome Probabilities

Outcome	$\frac{\partial x_i}{\partial t}$	$\frac{\partial x_i}{\partial v}$	$\frac{\partial x_i}{\partial p}$	$\frac{\partial x_i}{\partial c}$	$\frac{\partial x_i}{\partial q}$
Evasion/audit of high type (x_1)	−	−	−	+	−
Evasion/no audit of high type (x_2)	−	−	+ or −	+	−
Compliance by high type (x_3)	+	+	+	−	+
Audit of low type (x_4)	=	=	−	=	−
No audit of low type (x_5)	=	=	+	=	−

Note that,

$$\text{Prob}(\omega=Y|\lambda=1,\theta=u-v)=\frac{x_1}{x_1+x_2}=\frac{1}{1+p}, \tag{3}$$

as required by Proposition 1.

The effects of changes in tax rates, the amount of the loss, penalty rates, audit costs, and the proportion of taxpayers who can deduct their losses are shown in table 2.

Table 2 implies that an increase in the tax rate induces more compliance by high types but does not affect the IRS's audit strategy. A tax rate increase makes auditing more attractive for the IRS. High types respond to a rate increase by increasing the probability of compliance; otherwise, the IRS would change its decision rule so as to audit more low reports.[5] Taxpayer payoffs are linear functions of the tax rate, so a change in the rate will induce no change in the IRS's strategy. Thus an increase in the tax rate decreases detected and undetected tax evasion by high types, increases compliance by high types, and has no effect on the probability of a low type being audited. The effects of an increase in tax shelter losses are the same, and for the same reasons.

An increase in the penalty rate induces more compliance but less auditing. The increase makes auditing more attractive to the IRS, and so high types comply more. It also makes detected tax evasion more costly to taxpayers, so the IRS reduces the probability of a low report being audited. Thus an increase in the penalty rate decreases detected tax evasion, increases compliance by high types, and decreases the probability of a low type being audited. The effect on undetected tax evasion could be of either sign.

An increase in audit costs decreases the IRS's payoff but has no effect on taxpayer payoffs. This induces more evasion by high types but does not affect the audit strategy. Thus an increase in audit costs increases detected and undetected tax evasion, decreases compliance by high types, and does not affect the probability of a low type being audited.

[5] Yitzhaki (1974), in an extension of Allingham and Sandmo (1972), derives a similar result, but for different reasons. In these papers, audit probabilities are exogenous; higher tax rates induce greater compliance in Yitzhaki's paper due to risk aversion on the part of taxpayers.

An increase in the proportion of high types can be thought of as a tax law change further limiting the availability of losses from tax shelter partnerships to individuals. Such an increase makes it harder for high types to successfully pool with low types. This induces less evasion by high types but does not affect the probability of a low type being audited. An increase in the proportion of high types decreases detected and undetected tax evasion and increases compliance. Fewer low types are audited because there are fewer low types overall and the audit strategy is unchanged.

These parameter values changed significantly during the 1980s. Generally speaking, tax rates decreased while penalty rates were raised. The passive activity limitations reduced the number of taxpayers who could deduct losses from tax shelters, while the elimination of tax preferences such as rapid depreciation reduced the size of losses from most tax shelter investments. In the context of the model, these tax changes caused t and v to decrease, and p and q to increase. Table 2 indicates these changes have countervailing effects on taxpayer behavior and the IRS's audit strategy.

The Deficit Reduction Act of 1984 enacted §6111 and §6112, which imposed registration and disclosure requirements for tax shelters. This information is used by the IRS in its efforts to distinguish legitimate deductions from other deductions (Lore and Garbis 1985). In the context of this model, this represents an increase in the parameter η. The effect of an increase in η on the probability of the outcomes x_1 through x_5 is derived in Proposition 2.

Proposition 2: An increase in η causes:

(1) the probability of a high type reporting low and being audited, x_1, to increase if $\eta < -\log p$ and decrease if $\eta > -\log p$;
(2) the probability of a high type reporting low and not being audited, x_2, to increase if $\eta < -\log p$ and decrease if $\eta > -\log p$;
(3) the probability of a high type reporting high, x_3, to decrease if $\eta < -\log p$ and increase if $\eta > -\log p$;
(4) the probability of a low type being audited, x_4, to decrease, and;
(5) the probability of a low type not being audited, x_5, to increase.

An increase in η increases a high type's propensity to engage in tax evasion when η is low and decreases it when η is high. To see why, note that an increase in η induces an equal increase in δ^*, which is necessary to keep the high types indifferent between reports. Because ϵ^* is chosen to make the IRS indifferent between auditing and not auditing when $z = \delta^*$, the effect of η on ϵ^* depends on whether the *pdf* of the signal generated by a low type at δ^* increases or decreases when η increases. When η is sufficiently low, the density increases, which induces an increase in ϵ^*; when η is sufficiently high, the density decreases, which induces a decrease in ϵ^*. Since the probability of auditing a high type who claims the deduction is constant in equilibrium, an increase in η increases detected and undetected tax evasion when η is low, and decreases them when η is high.

An increase in η unambiguously decreases the probability of a low type being audited. The probability that the IRS audits, given that it receives a low report from a low type, is equal to,

$$\frac{x_4}{x_4 + x_5}.$$

Substituting the values of x_4 and x_5 from equations (2d) and (2e) yields,

$$\mathrm{Prob}(\omega=\mathrm{Y}\,|\,\lambda=0,\theta=u-v)=\frac{1}{1+pe^{\eta}}, \tag{4}$$

which is decreasing in η.

III. Optimal Information Acquisition

Since η is a measure of how well the signal discriminates between low and high types, an improvement in the signal is defined as an increase in η. The signal is valuable to the IRS if it increases expected government revenues net of audit costs. Expected gross government revenue, R, equals,

$$R=x_1[t(u+pv)]+x_2[t(u-v)]+x_3(tu)+(x_4+x_5)[t(u-v)]. \tag{5}$$

Substituting the values of x_1-x_5 from equations (2a) through (2e) yields $R=t(u-v+vq)$, which is independent of η. Thus the value of the signal manifests itself solely in its ability to reduce audit costs. Total audit costs, C, equal $c(x_1+x_4)$. Substituting in x_1 and x_4 from equations (2a) and (2d) yields,

$$C=\frac{c(1-q)[tv(1+p)-c+e^{\eta}(tpv(1+p)+c)]}{(1+pe^{\eta})^2[tv(1+p)-c]}. \tag{6}$$

First consider the value of an improvement in the signal when the improvement itself is costless. Interestingly, a costless improvement does not necessarily make the IRS better off.

Proposition 3: A costless increase in η increases net government revenues if and only if,

$$\eta>\log\left\{\frac{c(1+2p)-tpv(1+p)}{p[tpv(1+p)+c]}\right\}.$$

Corollary: If $c<tpv(1+p)$, a costless increase in η increases net government revenues.

An increase in η decreases net government revenues when η is low and c is high. Proposition 2 shows that an increase in η induces fewer audits of low returns filed by low types, but more low returns filed by high types when $\eta<-\log p$. The second effect dominates the first when the cost per audit is sufficiently high. Thus when the audit cost is high, an improvement in the ability of the signal to distinguish between types will increase audit costs, while not generating any increase in expected gross revenues. Therefore an improvement in the audit technology, even if costless, may make the IRS worse off without making the taxpayer better off.[6] Note the IRS cannot avoid this result by disregarding the signal; if the taxpayer believed the IRS would not use the signal, the IRS would be better off using it.

Henceforth I assume $c<tpv(1+p)$, which by the Corollary to Proposition 3 ensures that a costless improvement is beneficial to the IRS. I further assume the cost of the

---

[6] See Levine and Ponssard (1977) for a discussion of when information can make a player worse off in a strategic game. Baiman (1975) makes a similar point in the context of a choice of internal accounting systems.

signal is $s(\eta)$, where $s(0)=0$, $s'(\eta)>0$, $s''(\eta)>0$, and $s'(0)=0$. Finally, I assume the IRS can condition its choice of η on the observable taxpayer attributes u and v. That is, the IRS can spend different amounts on information acquisition for different types of taxpayers. The IRS chooses η to maximize $R-C(\eta)-s(\eta)$. Since $R-C(\eta)$ is concave in η for η sufficiently large, these assumptions ensure that an unique interior optimum level of precision, η^*, exists.

Examining how the value of the signal is affected by the model's parameters yields insights as to when expenditures by the IRS on information acquisition will be the most valuable. The effects of the model's parameters on η^* are shown in Proposition 4.

Proposition 4: η^* is,

(1) decreasing in q;
(2) increasing in t if and only if $\eta^* < -\log p$;
(3) increasing in v if and only if $\eta^* < -\log p$; and
(4) increasing in c if $\eta^* > -\log p$.

The optimal investment in information acquisition is decreasing in the proportion of high types because an increase in high types makes it harder for high types to pool with low types. When the signal is imprecise (i.e., $\eta < -\log p$), the value of an improvement in the signal is increasing in the tax rate and the amount of the loss from the tax shelter. When the signal is sufficiently precise, the value of an improvement is decreasing in the tax rate, decreasing in the amount of the loss from the tax shelter, and increasing in the cost of an audit. Thus the optimal level of investment in information will not vary monotonically in tax rates, tax shelter losses, or audit costs.

IV. Conclusions

This paper examines a model of taxpayer/IRS interaction. The effects and value of a signal regarding the taxpayer's type are analyzed. Four insights arise from the analysis. First, an improvement in the IRS's ability to detect tax evasion can induce more tax evasion in equilibrium. This result is driven by the fact that each player anticipates the other's behavioral response. Empirical studies of taxpayer compliance that do not consider these effects may lead to erroneous conclusions. Second, given that some level of compliance is achieved, expected gross government revenues do not depend on the precision of the signal. Third, the quality of the signal affects expected audit costs, but this effect need not be positive. This implies that information about taxpayers may have negative value to the IRS. Finally, the optimal level of investment in information acquisition by the IRS is not monotone in tax rates, the proportion of high types, the magnitude of tax shelter losses, or audit costs.

While this analysis is done in the context of the deductibility of losses from tax shelter partnerships, it is applicable to any tax compliance issue in which the important tax issue features a binary choice. It is applicable to settings such as whether a stock redemption should be treated as a dividend or stock sale, whether a corporation is eligible to be treated as an S corporation, whether a transaction is with a related party, and whether a corporate reorganization is taxable or tax-free. It may not be applicable to situations in which the tax issue is continuous, for example, whether a taxpayer overstates the level of charitable contributions.[7]

---

[7] See, however, Erard and Feinstein (1992), in which all taxpayers with a given level of income either report their true income or underreport by the same amount.

Appendix

Proof of Proposition 1

An examination of table 1 indicates that reporting $\theta = u$ when $\lambda = 0$ is a dominated strategy, since the taxpayer can do no better than paying $t(u-v)$. When $\lambda = 1$, the taxpayer prefers to report $\theta = u$ when $\omega = Y$ and $\theta = u - v$ when $\omega = N$. Similarly, choosing $\omega = Y$ when $\theta = u$ is a dominated strategy for the IRS. When $\theta = u - v$, the IRS prefers $\omega = Y$ if $\lambda = 1$ and $\omega = N$ if $\lambda = 0$. When $\theta = u - v$, the IRS chooses δ to make the high-income taxpayer indifferent between reporting high and reporting low. Thus δ^* solves:

$$\text{Prob}(\omega = Y \mid \lambda = 1)tv(1+p) = tv,$$

$$\frac{1}{1+e^{\delta-\eta}}tv(1+p) = tv, \text{ and}$$

$$\delta^* = \eta + \log p. \tag{A1}$$

A high type anticipates δ^* and chooses ϵ to make the IRS indifferent between auditing and not auditing when $z = \delta^*$. Thus ϵ^* solves:

$$\frac{c(1-q)pe^{\eta}}{(1+pe^{\eta})^2} = \frac{q\epsilon^* p[tv(1+p)-c]}{(1+p)^2}$$

$$\epsilon^* = \frac{e^{\eta}(1+p)^2 c(1-q)}{(1+pe^{\eta})^2 q[tv(1+p)-c]}. \tag{A2}$$

The IRS is indifferent between auditing and not auditing when $z = \delta^*$. It strictly prefers to audit when $z > \delta^*$ and strictly prefers to not audit when $z < \delta^*$ because the logistic distribution exhibits MLRP. The bounds on c ensure that $0 < \epsilon^* < 1$. Q.E.D.

Proof of Proposition 2

Differentiating equations (2a) through (2e) with respect to η yields the results. Q.E.D.

Proof of Proposition 3

Differentiating $(R-C)$ with respect to η yields,

$$\frac{ce^{\eta}(1-q)[(1+pe^{\eta})tpv(1+p)-c(1+2p-pe^{\eta})]}{(1+pe^{\eta})^3[tv(1+p)-c]} \tag{A3}$$

which is positive if and only if,

$$\eta > \log\left\{\frac{c(1+2p)-tpv(1+p)}{p[tpv(1+p)+c]}\right\}. \qquad \text{Q.E.D.}$$

Proof of Corollary

If $c < tpv(1+p)$, equation (A3) is positive for all η. Q.E.D.

Proof of Proposition 4

(1) Equation (A3) is the marginal benefit to the IRS from a costless increase in η. Differentiating equation (A3) with respect to q yields,

$$\frac{-ce^{\eta}\{pe^{\eta}[c+tpv(1+p)]+tpv(1+p)-c(1+2p)\}}{(1+pe^{\eta})^3[tv(1+p)-c]}, \tag{A4}$$

which is negative for all η when $c < tpv(1+p)$.

(2) Differentiating equation (A3) with respect to t yields,

$$\frac{vc^2e^{\eta}(1-q)(1+p)^2(1-pe^{\eta})}{(1+pe^{\eta})^3[tv(1+p)-c]^2}, \tag{A5}$$

which is negative if and only if $\eta > -\log p$.

(3) Differentiating equation (A3) with respect to v yields,

$$\frac{tc^2e^{\eta}(1-q)(1+p)^2(1-pe^{\eta})}{(1+pe^{\eta})^3[tv(1+p)-c]^2}, \tag{A6}$$

which is negative if and only if $\eta > -\log p$.

(4) Differentiating equation (A3) with respect to c yields,

$$\frac{e^{\eta}(1-q)\{pt^2v^2(1+p)^2(1+pe^{\eta})-c[2tv(1+p)-c](1+2p-pe^{\eta})\}}{(1+pe^{\eta})^3[tv(1+p)-c]^2}. \tag{A7}$$

The denominator of equation (A7) is positive, and the numerator is increasing in η. It suffices to show that the numerator is positive when $\eta=-\log p$. Substituting $\eta=-\log p$ into equation (A7) yields the result. Q.E.D.

References

Allingham, M. G., and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1 (November): 323–38.

Baiman, S. 1975. The evaluation and choice of internal information systems within a multiperson world. *Journal of Accounting Research* 13 (Spring): 1–15.

Beck, P., and W. Jung. 1989. Taxpayers' reporting decisions and auditing under information asymmetry. *The Accounting Review* 64 (July): 468–87.

Erard, B., and J. S. Feinstein. 1992. Honesty and evasion in the tax compliance game. Working paper.

Graetz, M., J. Reinganum, and L. Wilde. 1986. The tax compliance game: Toward an interactive theory of law enforcement. *Journal of Law, Economics, and Organization* 2 (Spring): 1–32.

Levine, P., and J. P. Ponssard. 1977. The value of information in some nonzero sum games. *International Journal of Game Theory* 6: 221–29.

Lore, P., and M. Garbis. 1985. Elements of IRS enforcement program outlined by Egger. *The Journal of Taxation* 63 (August): 121–22.

Reinganum, J., and L. Wilde. 1986. Equilibrium verification and reporting policies in a model of tax compliance. *International Economic Review* 27 (October): 739–60.

Shibano, T. 1990. Assessing audit risk from errors and irregularities. *Journal of Accounting Research* 28 (Supplement): 110–40.

Yitzhaki, S. 1974. A note on income tax evasion: A theoretical analysis. *Journal of Public Economics* 3 (May): 201–02.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 885-895

Lenders' Use of Accounting Information in the Oil and Gas Industry

Kwang-Hyun Chung
Pace University
Dimitrios Ghicas
Baruch College, City University of New York
Victor Pastena
University at Buffalo, SUNY

SYNOPSIS AND INTRODUCTION: A borrowing base is a line of credit set by the lender and secured by petroleum assets. Borrowers can draw on the borrowing base only to the extent that their investment opportunities justify the related interest expenses. This measure of the firm's capacity to obtain secured loans is footnoted in the long-term debt section of the annual reports.

In this paper we examine how lenders use accounting information to set the borrowing base of oil and gas firms. Determination of the borrowing base is a vital decision because it represents the lenders' exposure in the event that the borrower defaults.[1] We find evidence on lenders' use of accounting information by examining actual lending agreements as well as through tests of association.

Our sample consists of smaller petroleum firms that have a higher probability of default if unfavorable contingencies occur. The primary finding is that the value of firm's oil and gas reserves explain a large proportion of the variation in the firms' borrowing base and total outstanding debt.

---

[1] Lev and Ohlson (1982) advise that financial accounting research should not be limited exclusively to investor decisions but rather should include the broader spectrum of alternate applications of accounting information. Lenders' use of accounting seems to fit this broader application.

The authors appreciate the comments of Thomas Lechner, Larry Brown, Bob Hagerman, Jerry Han, David Jobson, Qin Luo, Gordon Richardson, Mike Rozeff, Tom Scott, and three anonymous referees. Of course, the authors are responsible for any remaining errors. Earlier versions of this article were presented at the University of Alberta, Pace University, SUNY at Buffalo, and the 1992 annual meeting of the American Accounting Association. The financial assistance of the CUNY Research Foundation is appreciated.

Submitted June 1992.
Accepted May 1993.

Unlike prior studies, we find that reserve recognition accounting (*RRA*) has a higher explanatory power than book values. Although major fluctuations in oil prices during the period of the study, 1984-1987, suggest that historical costs may be relatively poor indicators of changes in asset values, we observe that *RRA* information is used in setting the borrowing base for 21 of the 23 firms for which such agreements were available.

The evidence reported in this article complements prior research (Harris and Ohlson 1987, 1990; Ghicas and Pastena 1989)[2] and enhances the regulators' implication that their mandated *RRA* information is useful to users of financial statements.

Key Words: *Reserve recognition accounting, Lenders, Borrowing base, Debt contracts.*

Data Availability: *The data used in this study are available from the authors upon request.*

THE remainder of the article is organized as follows. In section I, we highlight the uses and limitations of book values and RRA and summarize the results of previous studies. In section II, we present the methodology to measure creditors' use of accounting information. Section III provides empirical results. In light of those results, we examine lending guidelines in the petroleum industry and lending agreements of petroleum firms in section IV, then summarize our conclusions in section V.

I. Background

Firms measure the book values of petroleum properties by successful efforts (SE) or full costs (FC). SE capitalizes costs for identifiable future benefits through discovery, acquisition, or development of specific oil and gas reserves. Costs that do not provide identifiable future benefits are expensed. FC capitalizes costs incurred in prospecting for, acquiring, and developing petroleum properties. Capitalized costs under FC are amortized over relatively large cost centers, usually operations over an entire country, as oil and gas are actually produced.[3]

When mandated by Congress to choose a uniform method for oil and gas accounting in the 1970s, the SEC chose to require RRA, but as supplemental or footnote information. The SEC does not support either SE or FC as the uniform accounting method for petroleum, in deference to the argument that there is little correlation between cost of exploration and value of discovery. In contrast, RRA attempts a dollar valuation based on the quantity of physical, proved reserves. The major RRA disclosure is an estimation of the present value of the net (of expenses) cash flows that will be generated by proved reserves. Proved reserves represent the oil that can be lifted on an economically competitive basis, given current prices and technology. The *RRA* present-value estimates are driven by estimates of the physical quantities of reserves and their lifting schedules.[4] With adjustments for taxes, SFAS No. 69 (FASB 1982) incorporates the SEC's *RRA* supplemental disclosures into generally accepted accounting principles (GAAP).

---

[2] These include Landsman (1986), Magliolo (1986), Ghicas and Pastena (1989), Harris and Ohlson (1987, 1990), and Shaw and Wier (1991). Magliolo (1986) also emphasizes profitability measures.

[3] Although Lilien and Pastena (1981, 1982) point out substantial intramethod differences with FC and SE, these are narrowed by initiatives of the Securities and Exchange Commission (SEC) in *Accounting Series Release No. 253* and No. 258. Moreover, the SEC now requires FC firms to apply an impairment test and write down realizable value when it is exceeded by book value. FC book values of individual petroleum properties often exceed the market values, but the SEC's "ceiling test" limits the aggregate carrying value of these properties to the present value of estimated future net revenues from proved reserves discounted by 10 percent. By means of this ceiling test, the SEC mandates RRA as the book value when FC book values exceed RRA.

The prior literature focuses on accounting information used by investors rather than lenders in the petroleum industry. Harris and Ohlson (1987) report that historical costs are important in explaining the market prices of petroleum firms, but present some evidence that *RRA* provides incremental explanation. Similarly, Shaw and Wier (1993) report that book values explain the market values of petroleum firms in the years 1985–1988.

The literature on accounting choice emphasizes book values over *RRA*. Implicitly, prior research assumes that book debt-to-equity ratios affect firms' borrowing capacity. Collins et al. (1981) use the debt-to-equity ratio and the proportion of private to total debt as a surrogate for contractual restrictions on petroleum firms. Similarly, Lilien and Pastena (1982) and Lys (1984) link book values to firms' financial contracts.

There is mixed evidence on the use of *RRA* in petroleum firms' contracts. Foster (1980) illustrates that non-GAAP accounting, including *RRA*, is utilized in petroleum contracting, and cites Juniper Petroleum and Houston Oil & Minerals as specific examples. In contrast, Frost and Bernard (1989) analyze the lending agreements of FC firms affected by the SEC's write-down requirements and found 15 agreements that use GAAP or modified GAAP in setting restrictions. Only five of the 20 agreements have covenants based on non-GAAP accounting or *RRA*. Frost and Bernard do not find any evidence of negative economic consequences associated with mandated reductions in affected firms' GAAP-based book values. Similarly, Malmquist (1990) does not report the use of *RRA* in his small sample of leverage contracts in the petroleum industry.

II. Methodology for Tests of Association

A regression model tests the ability of historical cost information and *RRA* to explain petroleum firms' secured borrowing base and their actual debt outstanding. The two dependent variables of interest are firms' borrowing base and total debt outstanding, which includes all current and long-term liabilities. We use the book value of oil and gas assets and *RRA* as independent variables. To control for possible heteroscedasticity, we use year-end market values to scale debt, *RRA*, and historical cost variables.[5]

---

[4] Despite its relevance, several aspects of *RRA* are subject to criticism. Adkerson (1979), Connor (1979), and Cooper et al. (1979) point out that there is considerable subjectivity in estimating proved reserves. Beyond problems in estimating the physical quantity of proved reserves, *RRA* involves a series of production and financial assumptions. Firms must estimate a multiyear schedule for lifting their proved reserves. This production schedule may change radically as demand varies. Selling prices in the future are assumed to remain unchanged from current spot levels, an assumption that likely deviates from that of equity market participants. Costs and tax factors must be estimated for the uncertain future. Also, *RRA* mandates a standardized rate to discount expected future net cash flows from lifting the proved reserves. The discount rate is not risk-adjusted and is applied uniformly whether the reserves are located in the United States or in riskier locations such as Angola or China. An alternate criticism of *RRA* is that the focus on proved reserves is actually too narrow.

[5] We use equivalent barrels of proved reserves of oil and gas as a scaler in the tests of robustness and sensitivity. As in prior research, 6,000 cubic feet of natural gas is considered the equivalent of a barrel of oil. Models are also tested with 8,000 cubic feet of natural gas as the conversion factor to determine the sensitivity of the previous assumption of equivalency.

The following sets of equations determine the ability of book values (model I) and *RRA* (model II) to explain the borrowing base (*BB*) and outstanding debt (*DOS*), respectively. The joint explanatory ability of the independent variables is determined in model III. For model I,

$$BB = a_0 + a_1 SEM^*BVOG + a_2 FCM^*BVOG + e, \text{ and} \tag{1}$$

$$DOS = b_0 + b_1 SEM^*BVOG + b_2 FCM^*BVOG + e. \tag{2}$$

For model II,

$$BB = a_0 + a_1 RRA + e, \text{ and} \tag{3}$$

$$DOS = b_0 + b_1 RRA + e. \tag{4}$$

For model III,

$$BB = a_0 + a_1 SEM^*BVOG + a_2 FCM^*BVOG + a_3 RRA + e, \text{ and} \tag{5}$$

$$DOS = b_0 + b_1 SEM^*BVOG + b_2 FCM^*BVOG + b_3 RRA + e. \tag{6}$$

Where *BB* = firms' stated borrowing base secured by their petroleum assets as given in the annual report; *DOS* = total debt outstanding, including all long-term and current liabilities; *SEM* = 1 if the firm uses the SE method, and 0 otherwise; *FCM* = 1 if the firm uses the FC method, and 0 otherwise; *BVOG* = the book value of the oil and gas assets of the firm; *RRA* = the present value of the net cash flows associated with proved reserves, as reported in conformity with SFAS No. 69; and e = the least squares regression error term.

The variables *BB*, *DOS*, *BVOG*, and *RRA* are scaled by market values. In equations (1), (2), (5), and (6), *BVOG* is used along with an interactive dummy variable that signifies the use of SE or FC. We expect both independent variables (book values and *RRA*) to be positively correlated with the borrowing base. As in Harris and Ohlson (1987), the RRA variable is the present value of expected net cash flows and represents anticipated revenues net of production costs and taxes discounted for the time value of money at 10 percent (consistent with SFAS No. 69).

To test whether *RRA* adds incremental explanatory power beyond that supplied by historical cost information, we compare the explanatory ability of the equations in the full model III with that of model I for book value. Similarly, to test whether SE and FC book values add incremental explanatory power beyond that supplied by *RRA*, we compare the explanatory ability of the full model III with that of model II for *RRA*. We use an F-test to assess the explanatory ability of the full model over those that use only book values or *RRA*.[6]

III. Sample and Regression Results

The sample period is the four years beginning in January of 1984 and ending in December of 1987. This was a period of fluctuating petroleum prices; the average spot

---

[6] $$F_{q,n-k} = \frac{(SSE_r - SSE_f)/q}{SSE_f/(n-k)},$$

where SSE_f = sum of squared prediction errors for the full model, SSE_r = sum of squared prediction errors for the reduced model, n = total number of observations, k = total number of estimated coefficients, and q = additional estimated coefficients in the full model over the reduced model.

price per barrel fluctuated from a high of $29 in 1984 to a low of approximately $12 in 1986. Prices had recovered to the $18 range by 1987. The sample consists of both listed and over-the-counter (OTC) firms but is restricted to firms that reported borrowing base and *RRA* information for at least two of the four years. The source of sample firms is COMPUSTAT Industry No. 1311, oil and gas explorer-producers.[7]

Determinants of the Secured Borrowing Base

We report the regression results tests explaining the borrowing base in table 1.[8] Model I (book value) has the lowest R^2 values for each year in all four years, ranging from 0.136 to 0.368; in seven out of eight cases, the coefficients of the book value variables are statistically significant in explaining firms' borrowing base. In model II, the coefficient of *RRA* is statistically significant for all four years, and R^2 values range from 0.507 to 0.595. The explanatory ability of model III, with R^2 values ranging from 0.48 to 0.59, does not uniformly dominate that of model II, which indicates that book values have limited explanatory ability.

In both models II and III, the coefficient of *RRA* is positive ($p<0.01$) for all four years. However, in model III the coefficients of book values are not positive at statistically significant levels in any of the four years for either FC or SE methods. An analysis of White's (1980) *t*-statistic, which accounts for the possible presence of heteroscedasticity, leads to the conclusion that *RRA* has explanatory ability while book values do not, if both are used as independent variables. Furthermore, an *F*-test indicates that the incremental contribution of *RRA* in the full model over model I for book value is statistically significant for all four years. The test also indicates that the incremental contribution of the book values over *RRA* is not statistically significant for any of the four years, which is consistent with the conclusion that *RRA* best explains the borrowing base.

Determinants of Total Debt

Regression results for total debt are presented in table 2. As in the case of the determinants of the borrowing base, model I for book value has the lowest R^2, ranging from 0.129 to 0.669, and the regression coefficients of the book value variables are generally statistically significant. Model II for *RRA* has greater explanatory ability, with the R^2 varying from 0.545 to 0.804. The full model III has R^2 values similar to the *RRA* model. In the full model, the coefficients of the book value are positive and significant in only one of the four years, but the coefficients of *RRA* are always statistically significant. As in the case of the borrowing base, White's *t*-statistics indicate that *RRA* has explanatory ability while book values do not.

Also, the *F*-test indicates that the incremental contribution of *RRA* over book value is statistically significant for all four years. With the exception of 1985, the statistical significance of the *F*-value for the incremental contribution of the book values is not

---

[7] The yearly sample available when total debt is the dependent variable varies between 54 and 74 firms. The yearly sample available when the borrowing base is the dependent variable varies between 40 and 56 firms. Regression tests exclude outliers with Cook's D-statistic in excess of 2. Details and descriptive information about sample firms are available from the authors.

[8] An analysis of Pearson correlations and collinearity diagnostics indicates some linear dependence among the proposed independent variables. Results must be interpreted in light of that linear dependency. Correlations among the independent variables are similar over the four years.

Table 1
Determinants of Borrowing Base/Market Value

	1984			*1985*			*1986*			*1987*		
	Model I (*n*=42)	*Model II* (*n*=45)	*Model III* (*n*=40)	*Model I* (*n*=52)	*Model II* (*n*=53)	*Model III* (*n*=52)	*Model I* (*n*=57)	*Model II* (*n*=56)	*Model III* (*n*=56)	*Model I* (*n*=52)	*Model II* (*n*=50)	*Model III* (*n*=50)
Adjusted R^2	0.348	0.586	0.589	0.136	0.507	0.481	0.368	0.595	0.596	0.168	0.547	0.536
F-value	11.920	63.270	19.660	5.020	54.550	16.760	17.290	81.860	28.060	6.130	60.260	19.830
Intercept	0.281	0.033	0.013	0.261	0.095	0.092	0.236	−0.043	0.014	0.498	0.149	0.181
White's *t*-statistic	2.330	0.310	0.140	1.270	0.780	0.750	1.090	−0.340	0.100	2.330	1.360	1.460
SE BV	0.213	–	−0.229	0.482	–	0.039	0.402	–	−0.244	0.304	–	−0.073
White's *t*-statistic	1.540	–	−1.310	3.650	–	0.210	3.750	–	−1.210	2.400	–	−0.410
FC BV	0.248	–	−0.185	0.356	–	−0.007	0.522	–	−0.169	0.286	–	−0.104
White's *t*-statistic	3.610	–	−1.370	2.580	–	−0.040	3.920	–	−0.760	2.320	–	−0.590
RRA	–	0.401	0.606	–	0.386	0.381	–	0.629	0.787	–	0.398	0.463
White's *t*-statistic	–	4.550	3.920	–	5.260	3.190	–	7.330	4.070	–	5.350	3.200
F-statistic for marginal contribution of *BV*	–	–	2.150	–	–	0.080	–	–	1.070	–	–	0.390
F-statistic for marginal contribution of *RRA*	–	–	24.170	–	–	33.650	–	–	31.190	–	–	24.206

Notes: Model I: $BB = a_0 + a_1 SEM{*}BVOG + a_2 FCM{*}BVOG + e$,
Model II: $BB = a_0 + a_1 RRA + e$, and
Model III: $BB = a_0 + a_1 SEM{*}BVOG + a_2 FCM{*}BVOG + a_3 RRA + e$,

where BB = the borrowing base; $SEM = 1$ if the firm uses the successful efforts (*SE*) method, and 0 otherwise; $FCM = 1$ if the firm uses the full cost (*FC*) method, and 0 otherwise; *BVOG* = the book value of the oil and gas assets of the firm; and *RRA* = the present value of oil and gas reserves.

Table 2
Determinants of Total Debt/Market Value

	1984			1985			1986			1987		
	Model I (n=56)	*Model II* (n=62)	*Model III* (n=54)	*Model I* (n=69)	*Model II* (n=70)	*Model III* (n=69)	*Model I* (n=75)	*Model II* (n=74)	*Model III* (n=74)	*Model I* (n=71)	*Model II* (n=68)	*Model III* (n=68)
Adjusted R^2	0.554	0.804	0.828	0.669	0.740	0.765	0.594	0.754	0.757	0.129	0.545	0.572
F-value	35.170	250.800	85.790	69.850	197.000	75.030	55.230	224.700	76.610	6.160	81.300	30.870
Intercept	−0.230	−0.610	−0.642	−0.245	−0.288	−0.361	−0.020	−0.337	−0.367	0.752	−0.072	0.044
White's *t*-statistic	−0.770	−2.690	−3.620	−1.540	−2.360	−3.050	−0.090	−1.880	−1.940	1.920	−0.330	0.260
SE BV	1.903	–	0.315	1.272	–	0.552	0.897	–	0.137	0.292	–	−0.467
White's *t*-statistic	2.730	–	1.590	6.980	–	2.730	3.080	–	0.590	1.490	–	−2.020
FC BV	0.862	–	−0.046	0.927	–	0.333	1.136	–	0.251	0.719	–	−0.167
White's *t*-statistic	8.700	–	−0.240	6.500	–	1.960	6.150	–	1.360	2.390	–	−0.600
RRA	–	1.244	1.222	–	0.911	0.600	–	1.231	1.039	–	0.892	1.068
White's *t*-statistic	–	6.950	4.620	–	9.570	5.460	–	8.940	5.410	–	4.790	4.220
F-statistic for marginal contribution of *BV*	–	–	5.400	–	–	6.290	–	–	1.380	–	–	3.080
F-statistic for marginal contribution of *RRA*	–	–	64.481	–	–	28.100	–	–	47.748	–	–	40.704

Notes: Model I: $DOS = b_0 + b_1 SEM{*}BVOG + b_2 FCM{*}BVOG + e$,
Model II: $DOS = b_0 + b_1 RRA + e$, and
Model III: $DOS = b_0 + b_1 SEM{*}BVOG + b_2 FCM{*}BVOG + b_3 RRA + e$,

where *DOS* = debt outstanding; *SEM* = 1 if the firm uses the successful efforts (*SE*) method, and 0 otherwise; *FCM* = 1 if the firm uses the full cost (*FC*) method, and 0 otherwise; *BVOG* = the book value of the oil and gas assets of the firm; and *RRA* = the present value of oil and gas reserves.

meaningful. This supports the conclusion that *RRA* plays a dominant role in explaining the secured borrowing base and the total outstanding debt of petroleum firms.[9]

IV. Lending Guidelines and Debt Contracts of Sample Firms

In light of this strong statistical evidence that the *RRA* is important information in lending decisions, we searched for corroborating evidence in banking industry guides and lending agreements.

The failures of Penn Square Bank and other lending institutions in the Southwest highlight the need for bankers to obtain proper collateral for their energy loans. Banking industry experts, including Busey and Simpson (1985), Webster (1987), and Sorrells (1987), advise bankers to focus on the present value of oil and gas reserves in setting the borrowing base associated with collateralized loans. They propose that lenders (1) obtain an oil and gas mortgage that meets the requirements of a Uniform Commercial Code financing statement and (2) limit the extension of credit to 50 percent of the present value of expected future cash flows from oil and gas properties.

Moreover, the lending literature explicitly questions the utility of book values. Sorrells (1987, 24) questions the applicability of GAAP's cost-based "values" of oil and gas properties for lending decisions: "Under the cost principle of accounting, the historic capitalized costs incurred in drilling and acquiring oil and gas wells, less accumulated depreciation, depletion, and amortization, are reported. This amount provides no indication of what future cash flows should be available for repayment of a loan."

The results of our search for the use of *RRA* in setting firms' borrowing base are summarized in table 3. We searched for lending information in firms' Form 10Ks, their public and private lending agreements, and their write-ups in Moody's. We restricted the search for evidence of firms' use of *RRA* in setting borrowing base or restricting managerial actions to the 33 firms listed on either the New York or American Exchanges because the Form 10Ks of listed firms are more readily available. Firms often include lending agreements as exhibits to their Form 10K.

Ten of the 33 firms had no agreements reported as amendments to their Form 10Ks. We consider the evidence for these 10 firms as indeterminate. Accordingly, our conclusions on the use of *RRA* in lending agreements are based on the remaining 23 firms. For these, we found 21 uses of *RRA*. Given the preponderance of *RRA* in the agreements of listed firms, we have no reason to believe that the agreements of the OTC firms are written differently. This evidence indicates substantial use of *RRA* in lending agreements and is at variance with earlier evidence by Frost and Bernard (1989) and Malmquist (1990).[10] However, the use of *RRA* in lending agreements is consistent with the reported ability of *RRA* to explain firms' borrowing base.

---

[9] The reported results use petroleum firms' year-end market values as a scaler for *RRA* present values and the book values with FC and SE. This controls for possible heteroscedasticity. If we use the unscaled book values and unscaled *RRA* as independent variables, *RRA* continues to be a statistically significant explanatory variable for both borrowing base and debt outstanding, while book values are not. Similarly, when we add the profitability of oil and gas operations as defined in Harris and Ohlson (1987) to the full models, the coefficients of *RRA* continue to be statistically significant with the expected sign, while the coefficients of the other proposed explanatory variables are not.

[10] Frost and Bernard (1989) searched for use of book value or *RRA* in restrictive covenants only. In contrast, we searched for references to *RRA* in either the covenants or the definition of the borrowing base. Some agreements use *RRA* to set the amount that can be borrowed but then include restrictive covenants based on book values. This type of agreement would have been classified as based on book value by Frost and Bernard but is considered *RRA*-based in the current study.

Table 3
Utilization of Reserve Recognition Accounting (RRA)
by Sample Firms

Company	*RRA Used?*	*Source of RRA Information*
Adobe Resources	Yes	Private agreement
Alamco Inc.	Yes	10K, Moodys
Altex Industries	NA	
Apache Corp.	Yes	Private agreement
Barnwell Industries	NA	
Baruch Foster	No	
Belden & Blake	Yes	Moodys, private agreement
Convest Exploration	Yes	10K, private agreement
Damson Oil	Yes	Moodys, private agreement
Freeport McMoran	Yes	10K, private agreement
Global National Resources	NA	
Hershey Oil	NA	
Kaneb Services	NA	
Louisiana Light & Electric	Yes	10K
MCO	Yes	Moodys, private agreement
Mesa	NA	
Mitchell Energy	Yes	Moodys, public agreement
Noble Affiliates	Yes	Moodys, private agreement
NRM Energy	Yes	Moodys, private agreement
Ocean Drilling & Exploration	No	
Patrick Petroleum	NA	
Plains Petroleum	Yes	10K, private agreement
Pogo Producing	Yes	Moodys, private agreement
Presidio Oil	Yes	Moodys, private agreement
Sage Energy—LP	NA	
Samson Energy—LP	Yes	10K, private agreement
Summit Energy	NA	
Sunshine Mining	Yes	Private agreement
Transco Expl—LP	Yes	10K, private agreement
Triton	Yes	Moodys, private agreement
Wainoco Oil	Yes	Moodys, private agreement
Wilshire Oil TX	Yes	Moodys, private agreement
Wichita River	NA	

Note: NA indicates that a lending agreement was not available. All 33 sample firms were listed on either the New York or American Exchange.

V. Conclusion

This study finds that *RRA* is the pivotal information used by lenders in setting firms' borrowing base. Industry publications advise lenders to consider *RRA*, *not* book values, in setting lending limits. We show that most listed petroleum firms have lending agreements that stress *RRA* in setting the borrowing base. Moreover, our multiple regression results indicate that *RRA* is highly associated with both borrowing base and total debt outstanding. The empirical tests do not indicate that book values are used in setting borrowing limits or in explaining the total debt outstanding. The significance of *RRA* and the low incremental explanatory ability of book values holds whether petroleum prices are rising, falling, or stable.

References

Adkerson, R. 1979. Can reserve recognition accounting work? *Journal of Accountancy* 148 (September): 72–81.

Beaver, B., A. Christie, and P. Griffin. 1980. The informational content of SEC Accounting Series Release No. 190. *The Journal of Accounting & Economics* 2 (June): 127–58.

Belsley, D., S. Kuh, and R. Welsh. 1980. *Regression Diagnostics*. New York: John Wiley.

Busey, P. G., and W. G. Simpson. 1985. The documentation and structure of oil and gas loans. *The Journal of Commercial Bank Lending* 67 (January): 33–41.

Collins, D., M. Rozeff, and D. Dhaliwal. 1981. The economic determinants of the market reaction to proposed mandatory accounting changes in the oil and gas industry: A cross-sectional analysis. *Journal of Accounting & Economics* 3 (March): 37–71.

Connor, J. 1979. Reserve recognition accounting: Fact or fiction. *Journal of Accountancy* 148 (September): 92–99.

Cooper, K., S. Flory, S. D. Grossman, and J. C. Groth. 1979. Reserve recognition accounting: A proposed disclosure framework. *Journal of Accountancy* 148 (January): 82–91.

Financial Accounting Standards Board. 1977. *Statement of Financial Accounting Standard No. 19: Financial Accounting and Reporting by Oil and Gas Producing Companies*. Stamford, CT: FASB.

———. 1982. *Statement of Financial Accounting Standard No. 69: Disclosures about Oil and Gas Producing Activities*. Stamford, CT: FASB.

Foster, G. 1980. Accounting policy decisions and capital market research. *Journal of Accounting & Economics* 2 (March): 29–62.

Frost, C. A., and V. L. Bernard. 1989. The role of debt covenants in assessing the economic consequences of limiting capitalization of exploration costs. *The Accounting Review* 64 (October): 788–808.

Ghicas, D., and V. Pastena. 1989. The acquisition value of oil and gas firms: The role of historical costs, reserve recognition accounting, and analysts' appraisals. *Contemporary Accounting Research* 6 (Fall): 125–42.

Harris, T., and J. Ohlson. 1987. Accounting disclosures and the market's valuation of oil and gas properties. *The Accounting Review* 62 (October): 651–70.

———, and ———. 1990. Accounting disclosures and the market's valuation of oil and gas properties: Evaluation of market efficiency and functional fixation. *The Accounting Review* 65 (October): 764–80.

Landsman, W. 1986. An empirical investigation of pension and property rights. *The Accounting Review* 61 (October): 962–91.

Lev, B., and J. Ohlson. 1982. Market-based empirical research in accounting: A review, interpretation, and extension. *Journal of Accounting Research* 20 (Supplement): 249–322.

Lilien, S., and V. Pastena. 1981. Intramethod comparability: The case of the oil and gas industry. *The Accounting Review* 56 (October): 690–703.

———, and ———. 1982. Determinants of intra-method choice in the oil and gas industry. *Journal of Accounting & Economics* 4 (December): 145–70.

Lys, T. 1984. Mandated accounting changes and debt covenants: The case of oil and gas accounting. *Journal of Accounting & Economics* 6 (March): 39–65.

Malmquist, D. H. 1990. Efficient contracting and the choice of accounting method in the oil and gas industry. *Journal of Accounting & Economics* 12 (January): 173–206.

Magliolo, J. 1986. Capital market analysis of reserve recognition accounting. Supplement to *Journal of Accounting Research* 24: 69–111.

Securities and Exchange Commission. 1978a. *Accounting Series Release No. 253: Adoption of Requirements for Financial Accounting and Reporting Practices for Oil and Gas Producing Activities.* Washington, D.C.: U.S. Government Printing Office.

——. 1978b. *Accounting Series Release No. 258: Oil and Gas Producers—Full Cost Accounting Practices.* Washington, D.C.: U.S. Government Printing Office.

Shaw, W., and H. Wier. 1993. Organizational form choice and the valuation of oil and gas producers. *The Accounting Review* 68 (July): 657–67.

Sorrells, J. G. 1987. Coping with problem oil and gas credits: A review of the basic factors. *The Journal of Commercial Bank Lending* 69 (May): 24–30.

Webster, F. T. 1987. Coping with problem oil and gas credits: Overview. *The Journal of Commercial Bank Lending* 69 (May): 23–24.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48: 817–38.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 896–912

The Stock Price Effects of Alternative Types of Management Earnings Forecasts

Grace Pownall
Emory University
Charles Wasley
Washington University
Gregory Waymire
Emory University

SYNOPSIS AND INTRODUCTION: This paper examines the stock price effects of alternative types of management earnings forecasts. Beyond deciding whether to disclose forecasts, managers must decide whether to issue a point projection or a more qualitative estimate (e.g., a bounded range), and whether to project interim or annual earnings or both. Our empirical tests assess differences in the information content of management earnings forecasts that differ by form and horizon.

Our tests provide a comprehensive investigation of the price effects of these alternative forecast disclosure types. While an extensive literature exists on the relation between management forecasts and stock prices, most previous studies examine only point and range forecasts of annual earnings (e.g., Penman 1980; Ajinkya and Gift 1984; Waymire 1984; McNichols 1989; Pownall and Waymire 1989). Exceptions include Lev and Penman (1990), Patell (1976), and Baginski et al. (1993). Lev and Penman (1990) include lower and upper bound forecasts for part of their sample period, but do not examine these disclosure forms separately. Patell (1976) provides evidence on mean price changes associated with a pooled sample of annual minimum and maximum forecasts. Baginski et al. (1993) examine alternative forecast forms.

Prior analyses of managers' disclosure incentives speculate that investors may condition their assessment of forecast information on disclosure

We appreciate the financial support of the John M. Olin School of Business at Washington University, the research assistance of Charles Kile, and the programming of Ronald Harris. We also appreciate comments and suggestions from Steve Baginski, Nick Dopuch, Carol Frost, Mahendra Gupta, Ron King, Mort Pincus, seminar participants at Emory University, two anonymous reviewers, and the associate editor.

Submitted October 1991.
Accepted April 1993.

form and horizon. For instance, King et al. (1990) suggest that forecast disclosures emerge as voluntary managerial actions to reduce costly information asymmetry in capital markets. Under the "expectations adjustment" hypothesis, managers have incentives to acquire and maintain a reputation for credible disclosure. Rational investors recognize that disclosure quality varies systematically by disclosure form and will discount qualitative projections or those issued with longer horizons.

Policy debates on mandatory disclosure of qualitative information, such as the recent SEC debates over the content of "Management Discussion and Analysis" disclosures, and deliberations on forecast disclosure in the 1970s (see King et al. 1990), also suggest a need for evidence on the information content of qualitative prospective disclosures and alternative forms of forecasts.[1]

Our primary tests are based on a sample of 1,252 forecasts disclosed by 91 firms between July 1, 1979 and December 31, 1987. Several conclusions emerge from these tests. First, forecast disclosures remain highly informative even when including other disclosure types not analyzed in prior studies. Second, forecasts are less informative than earnings announcements for our full sample, a finding that is inconsistent with earlier results in Pownall and Waymire (1989). Third, differences across forecast forms are not significant at conventional levels. Fourth, interim forecasts are significantly more informative than annual projections. This result is driven largely by maximum forecasts, which are highly informative and more frequent in the interim forecast subsample.

We document several additional regularities that may be of interest to researchers. First, point and range annual forecasts comprise less than 20 percent of our sample. This suggests that the incidence of voluntary management forecast disclosure is possibly far greater than suggested by previous studies. Second, range forecasts tend to be quite inaccurate *ex post*. Actual earnings per share (EPS) fell outside the forecasted bounds in more than 50 percent of our range forecasts. Third, forecasts that are more qualitative tend to be issued over longer horizons. Minimum forecasts are issued over the longest horizons for our sample, and interim point projections have the shortest horizons. Finally, extensions to our primary tests provide some evidence that maximum forecasts have significant negative price effects, and that for point forecasts, forecast revisions are highly informative.

Key Words: *Earnings forecasts, Voluntary disclosure, Stock prices.*

Data Availability: *The disclosure sample used in this paper is available from the first author.*

---

[1] For evidence on the information content of non-forecast qualitative prospective disclosures, see Hoskins et al. (1986) and Thompson et al. (1987).

THIS paper is organized as follows. Sample selection and data characteristics are described in section I. The design and results of our primary stock price tests are provided in section II. Section III contains diagnostics on our design, and extensions to our primary tests. We summarize the results and offer conclusions in section IV.

I. Sample Selection and Description

In the first stage of sample selection, we identified management forecasts from randomly selected time periods stratified by calendar year and quarter. From the firms identified in stage one, we randomly chose a sample and collected all management forecasts released by these firms between July 1, 1979 and December 31, 1987. The purpose of this approach was to obtain a relatively representative sample at low cost.

Stage-One Sample

We first limited our sample period to 1980–1987, and selected a random sample of forecasts included in the *Dow Jones News Retrieval Service* (*DJNRS*) or *The Wall Street Journal* (*WSJ*), stratified across years and calendar quarters. For each year 1980–1983, we randomly selected a calendar quarter without replacement, and then reversed the order of selected quarters for the years 1984–1987. For each selected quarter, we randomly selected a full five-trading-day week. For each week, we collected all earnings forecasts from the *DJNRS* and the *WSJ* that were (1) attributed to the firm or its management, and (2) expressed as a point, range, minimum, or maximum forecast. This generated a stage-one sample of 444 management forecasts.

Panel A of table 1 describes sampling periods, and panel B describes the frequency and forecast horizons for the stage-one sample, where forecast horizon is the number of calendar days between forecast disclosure and the end of the period to which the forecast applies. Annual point and range forecasts comprise 19.6 percent of the sample suggesting that forecast disclosure frequency is possibly five times as great as indicated by samples based exclusively on point and range annual forecasts. This ratio is attributable to two factors. First, interim forecasts comprise 41.2 percent of the stage-one sample. Second, 69.4 percent of the forecasts are minimum or maximum projections.

The frequency of minimum and maximum forecasts is apparently higher in our sample than in prior periods. In the Patell (1976) sample of annual forecasts selected from *WSJ* during 1963–1968, 204 (60.7 percent) are point or range forecasts, and the remaining 132 (39.3 percent) are minimums or maximums. As a benchmark, 179 (69 percent) of our annual forecasts from the *DJNRS* were also reported in the *WSJ*.[2] Of these, 57 (31.8 percent) are points or ranges, and the remaining 122 (68.2 percent) are minimums or maximums.

The relative frequency of minimum and maximum forecasts varies across the interim and annual subsamples. For the interim sample, the frequency of minimum and maximum forecasts is equal. Annual minimums outnumber annual maximums by

---

[2] To assess how much difference is likely between samples collected from the *WSJ* and samples collected from the *DJNRS*, we checked both the *WSJ* and *The Wall Street Journal Index* (*WSJI*) for inclusion of the forecasts from our *DJNRS* search. We found that the loss of data from using the *WSJ* or the *WSJI* was substantial. Of the 444 forecasts in our stage-one sample, only 215 (48.4 percent) are available in the *WSJ* and 126 (28.4 percent) in the *WSJI*. In contrast, only 17 *WSJ* forecasts were not on the *DJNRS*.

Table 1
Descriptive Statistics for Stage-One Sample of Management Earnings Forecasts (n=444)

Panel A. Time Periods Used in Constructing Stage-One Sample:

Calendar Year	*Calendar Quarter*	*Dates Sampled*	*Number of Forecasts*
1980	1	2/25–2/29	56
1981	4	12/14–12/18	74
1982	2	4/12–4/16	66
1983	3	8/1–8/5	49
1984	3	7/9–7/13	37
1985	2	6/17–6/21	88
1986	4	10/20–10/24	56
1987	1	3/2–3/6	18

Panel B. Frequency of Alternative Forecasts and Summary Statistics on Forecast Horizon:[a]

Type of Forecast	*N*	*Percent of Sample*	*Mean Horizon*	*Median Horizon*
Interim Point	36	8.1	35.92	15.5
Interim Range	13	2.9	37.62	16
Interim Minimum	67	15.1	111.82	74
Interim Maximum	67	15.1	54.30	34
Annual Point	57	12.8	181.98	172
Annual Range	30	6.8	157.43	158.5
Annual Minimum	139	31.3	231.22	259
Annual Maximum	35	7.9	148.80	153

[a] Forecast Horizon is the number of calendar days between forecast disclosure and the end of the period to which the forecast applies. Panel B was reproduced as table 1 of King et al. (1990).

about 4:1. These results suggest a relation between forecast horizon and type of forecast. The evidence in panel B of table 1 also suggests horizons are systematically longer for minimums (relative to other types) in both the interim and annual subsamples. The mean forecast horizon for the interim minimums is 111.82 calendar days compared with 70.56 days for the full interim subsample. The mean forecast horizon is 231.22 days for the annual minimums while the mean for the full annual subsample is 200.93 days.

Stage-Two Sample

From our stage-one sample, we randomly selected 150 forecasts (by 148 firms). We required each firm to have *Value Line* analyst forecast data (loss of 57 firms) and CRSP daily stock returns data (either NYSE/AMEX or NASDAQ). These selection filters yielded a sample of 91 firms. For these firms, we performed a full-text search on the *DJNRS* between July 1, 1979 and December 31, 1987, using the keywords "earnings," "income," and "profits." We kept forecasts attributed to the firm or its management that were released on or before the final day of the applicable fiscal period. The final sample consists of 1,252 forecasts by 91 firms.[3] The number of forecasts per firm ranged from one to 59. Three (eight) firms issued only interim (annual) forecasts over the time period. Three firms released only point and range forecasts, and nine firms disclosed only minimum and maximum forecasts.

We collected quarterly earnings amounts and disclosure dates for these companies between July 1, 1979 and December 31, 1987. We collected *Value Line* forecasts to serve as proxies for investor earnings expectations before all management forecast and earnings announcements.[4] We measured the disclosure's unexpected component (UD_{ikt}) as:

$$UD_{ikt} = (MF_{ikt} - AF_{it}) / P_{it-5}, \tag{1}$$

where i indexes firms; k indexes disclosure forms and/or horizons; t indexes trading days; MF_{ikt} is the amount disclosed for points, minimums, and maximums, and is the midpoint for range forecasts; AF_{it} is the most recent *Value Line* forecast; and P_{it-5} is the firm's stock price five days before disclosure. A minimum or maximum forecast is assumed to correspond to the mean of the manager's expectations.[5] For earnings announcements, *UD* is defined analogously, except we replaced MF_{ikt} with actual EPS in the numerator of equation (1). We examined the mean forecast error (*FE*) through the following equation:

$$FE_{ikt} = (ACT_{it} - MF_{ikt}) / P_{it-5}, \tag{2}$$

where ACT_{it} is the actual EPS for the period to which the management forecast applies.

Several aspects of table 2 are noteworthy. First, minimum forecasts are associated with longer average *MF* horizons. Second, annual minimums are the most frequent forecasts. Third, *AF* horizons are fairly homogeneous—only the interim points have a mean *AF* horizon greater than 50 days. Therefore, our empirical tests do not control for the effects of varying horizons between forecast types and should be interpreted accordingly. Fourth, all forecast types, except annual maximums, have negative mean *UD*s and a percentage of negative *UD*s greater than 50 percent. Finally, the percentage

---

[3] This sample constitutes a lower bound on the number of management forecasts issued by the 91 sample firms during the time period. First, we collected all quantifiable forecasts but not more vague projections of future results that did not include a specific number or refer to a defined time period. Second, in our search we used only the keywords "earnings," "income," and "profits." A loss forecast might have been contained in a *DJNRS* document that did not use these words to refer either to the forecast or to past, current, or future information in the same press release. Not including "loss" in the list of keywords would result in failing to identify the observation. However, we believe the loss of observations and resulting bias in our inferences is minimal. We used an extensive list of keywords (available from the authors) to generate our stage-one sample, and found that a full-text search using only the keywords "earnings," "income," and "profits" would have identified all the articles containing our stage-one forecasts.

[4] Our choice of *Value Line* analyst forecast data was largely arbitrary. *Standard & Poor's* forecast data were not available, because we used quarterly earnings announcements and some interim forecasts. We did not use IBES since quarterly data were not available for part of our sample period. See Philbrick and Ricks (1991) for a comparison of *Value Line* and IBES analyst forecast data.

[5] Our primary analysis assumes that *UD* is an appropriate proxy for the information communicated in minimum and maximum forecasts. We examine alternative empirical specifications for these disclosures in section III.

Table 2

Summary Statistics for Characteristics of Management Forecasts and Earnings Announcements

Type of Forecast	N	*Percent of Sample*	*Mean MF Horizon*[a]	*Mean AF Horizon*[b]	*Mean UD*[c]	N (%) *UD*<0	*Mean FE*[d]	N (%) *FE*<0
Management Forecasts	1252	100.0	142.6	45.3	−0.009	820 (65.5)	−0.011	541 (43.2)
Total Points	238	19.0	119.9	48.2	−0.009	134 (56.3)	−0.002	91 (38.2)
Total Ranges	168	13.4	128.4	44.7	−0.006	90 (53.6)	−0.013	87 (51.8)
Total Minimums	580	46.3	184.5	44.2	−0.015	461 (79.5)	0.005	122 (21.0)
Total Maximums	266	21.2	80.4	45.6	0.005	135 (50.8)	−0.054	241 (90.6)
Total Interim	452	36.1	50.4	44.1	−0.005	280 (61.9)	−0.010	204 (45.1)
Total Annual	800	63.9	194.7	46.0	−0.012	540 (67.5)	−0.012	337 (42.1)
Interim Points	90	7.2	26.3	52.0	−0.008	56 (62.2)	0.0004	21 (23.3)
Interim Ranges	56	4.5	44.0	42.4	−0.003	28 (50.0)	−0.003	27 (48.2)
Interim Minimums	154	12.3	69.7	41.5	−0.006	117 (76.0)	0.003	19 (12.3)
Interim Maximums	152	12.1	47.5	42.6	−0.002	79 (52.0)	−0.033	137 (90.1)
Annual Points	148	11.8	176.9	45.8	−0.009	78 (52.7)	−0.004	70 (47.3)
Annual Ranges	112	9.0	170.5	45.8	−0.008	62 (55.4)	−0.018	60 (53.6)
Annual Minimums	426	34.0	226.0	45.2	−0.018	344 (80.8)	0.006	103 (24.2)
Annual Maximums	114	9.1	124.4	49.6	0.013	56 (49.1)	−0.084	104 (91.2)
Earnings Announcements	3019	100.0	NA	41.4	−0.005	1427 (47.3)	NA	NA

[a] Management Forecast Horizon (MF Horizon) is the number of calendar days between management forecast disclosure and the final day of the period to which the management forecast applies. Medians are similar to means in all cases.

[b] Analyst Forecast Horizon (AF Horizon) is the number of calendar days between the analyst forecast date and the management forecast date. Medians are similar to means in all cases. Four observations (all annual point forecasts) are not included in the AF Horizon analysis because they were for EPS more than one year into the future and therefore could not be matched with *Value Line* analyst forecasts.

[c] *UD* (unexpected component of the disclosure)$=(MF_{ikt}-AF_{it})/P_{it-5}$, as in equation (1). Four observations (all annual point forecasts) are not included in the *UD* analysis because they were for EPS more than one year into the future and therefore could not be matched with *Value Line* analyst forecasts.

[d] *FE* (ex post forecast error)$=(ACT_{it}-MF_{ikt})/P_{it-5}$, as in equation (2).

Table 3
Summary Statistics for the Relation Between Management Forecasts and Actual EPS

Panel A. Absolute Difference Between Point Forecasts and Actual EPS Scaled by Price at t−5:

Point Forecasts	*N*	*Mean*	*Median*	*Range*	*Interquartile Range*
All	238	0.006	0.002	0.0 to 0.069	0.0007 to 0.006
Annual	148	0.008	0.003	0.0 to 0.069	0.0009 to 0.008
Interim	90	0.003	0.001	0.0 to 0.032	0.0004 to 0.004

Panel B. Frequency of EPS Within and Outside Bounds of Range Forecasts:

		Actual EPS		
Range Forecasts	*N*	*Below Range (%)*	*Within Range (%)*	*Above Range (%)*
All	168	50 (29.8)	76 (45.2)	42 (25.0)
Narrow Spreads[a]	84	27 (32.1)	32 (38.1)	25 (29.8)
Wide Spreads	84	23 (27.4)	44 (52.4)	17 (20.2)
Annual	112	39 (34.8)	44 (39.3)	29 (25.9)
Narrow Spreads	57	19 (33.3)	21 (36.8)	17 (20.2)
Wide Spreads	55	20 (36.4)	23 (41.8)	12 (21.8)
Interim	56	11 (19.6)	32 (57.1)	13 (23.2)
Narrow Spreads	28	4 (14.3)	15 (53.6)	9 (32.1)
Wide Spreads	28	7 (25.0)	17 (60.7)	4 (14.3)

[a] The variable spread is the upper bound of the range forecast minus the lower bound of the range forecast standardized by the midpoint of the range. Narrow spreads are those range forecasts for which the value of the spread is at or below the median, and wide spreads are those for which the value of the spread is above the median. We computed the median value of the variable spread separately for all ranges (mean = 0.164, median = 0.065), annual ranges (mean = 0.117, median = 0.048), and interim ranges (mean = 0.257, *median* = 0.148).

of negative forecast errors is relatively low (high) for the minimum (maximum) forecasts. If a minimum forecast is "accurate," then actual EPS should exceed the forecast, and vice versa for the maximum projections. For all minimums, 21 percent of actual EPS observations fall below the minimum, and for only 9.4 percent of the maximums does actual EPS fall above the maximum. This suggests that minimum and maximum forecasts are relatively accurate predictors of subsequent earnings.[6]

Table 3 provides summary statistics on the *ex post* accuracy of point and range forecasts. Panel A shows summary statistics on the absolute difference between forecasted and actual EPS scaled by price at $t-5$ for the point forecasts. The mean absolute scaled difference is 0.006 for the full sample of point forecasts, and is several times larger for annual forecasts than for interim projections.[7] Panel B describes the frequency with which actual EPS falls outside (either above or below) the bounds of sample range forecasts. Actual EPS is within the forecasted bounds in only 44.6 percent of the 168 ranges, and is more likely to be within the bounds of interim forecasts and forecasts with wider "spreads" (i.e., the difference between the lower and upper bounds). The inaccuracy of the range forecasts suggests caution in aggregating these forecasts in samples for empirical tests.

---

[6] We recomputed summary statistics in table 2 for subsamples including initial forecasts during the fiscal period, follow-up forecasts (i.e., forecast revisions), and final forecasts issued during the year. The summary statistics in table 2 are similar for subsamples, except that follow-up forecasts are typically annual projections and final forecasts are more likely to be point forecasts. We also recomputed the statistics excluding forecasts issued in 1982, since McNichols (1989) finds that forecasts issued in 1982 differ from forecasts issued in other years between 1979 and 1983. The only material difference in the forecast distribution is that there are slightly fewer maximums and slightly more minimums in 1982 than in the whole period.

II. Primary Tests

Design and Hypothesis Tests

For each sample firm, we specified the following general model of the return-generating process:

$$R_{it}=\alpha_i+\beta_i R_{mt}+\sum_{k=1}^{K}\gamma_{ik}UD_{ikt}+\epsilon_{it}, \tag{3}$$

where R_{it} = return for security i on day t;[8] R_{mt} = value-weighted market index return on day t; α_i, β_i = market model parameters; UD_{ikt} = amount defined by equation (1) on the trading day before and the date of the relevant management forecast or quarterly earnings announcement, and zero otherwise;[9] γ_{ik} = parameter measuring the response of the firm's stock price to alternative types of disclosures (K = total number of disclosure types released by firm i); and ϵ_{it} = random disturbances drawn from a multivariate normal distribution with an expected value of zero (for all firms) that are assumed to be independent and identically distributed through time for each firm, but that may be heteroscedastic and contemporaneously cross-correlated.

Equation (3) is a conditional market model similar to that used in Pownall and Waymire (1989). We estimated firm-specific coefficients to capture the stock price effects of alternative disclosure types. Including quarterly earnings announcements in the regression reduces problems caused by omitted variables, and provides a benchmark for evaluating the effects of forecasts.[10] We estimated several variants of equation (3) beginning with a simple specification that included a single forecast coefficient, and then adapting the model to allow for differing effects of alternative forms and horizons. This permits better comparison between our sample and those used in prior research.

---

[7] We also examined unscaled forecast errors. The mean absolute (unscaled) difference is 21 cents for the full sample of point forecasts, and is approximately four times larger for annual forecasts than for interim projections.

[8] We used raw returns to compute the results reported in this paper. We replicated these analyses using continuously compounded returns (i.e., the natural logarithm of one plus the raw return) with no impact on our conclusions.

[9] When a firm issued multiple forecasts of the same type on the same date, we used an average of the *UD*s. One firm released a single point forecast, and the *UD* for this forecast was zero. Because no coefficients could be estimated for this firm, we dropped it from the sample in the subsequent stock price tests.

[10] Waymire (1984) provides evidence that multiple forecasts are often released simultaneously and frequently disclosed with earnings announcements. Including earnings announcements along with multiple forecasts allows us to control for these effects. Diagnostic tests that examine the sensitivity of our results to other simultaneous disclosures are discussed in section IV.

We tested three hypotheses for each variant of equation (3). The first assesses the information content of the various forecast types; which is stated as:

$$H_{01}: \sum_{i=1}^{N} \gamma_{ik} = 0.$$

This test is in the spirit of traditional event studies where tests are performed on average abnormal returns or the coefficient from a cross-sectional regression of abnormal returns on unexpected earnings.[11]

The second null hypothesis is that each of the coefficients is jointly equal to zero:

$$H_{02}: \gamma_{ik} = 0 \qquad \forall i.$$

Based on prior research, we expect more individually significant coefficients to be positive. This test is more likely to reject the null than tests on the average coefficient if some but not all firms' forecasts are informative.

Our third hypothesis addresses average, sample-wide differences in the coefficients for different forecast types. In null form, the third hypothesis is that the average coefficient for one forecast type is equal to the average coefficient for another forecast type:

$$H_{03}: \bar{\gamma}_{ik} = \bar{\gamma}_{ij} \qquad k \neq j.$$

In the discussion to follow, we report the results of only those tests of the third hypothesis that enable us to reject the null.

We estimated the firm-specific models in equation (3) jointly with generalized least squares (GLS), following Zellner's (1962) seemingly-unrelated-regressions (SUR) approach. We test the hypotheses jointly across the sample as linear restrictions on coefficients, and use the covariance structure of the residuals to calculate the finite sample *F*-statistics presented in Schipper and Thompson (1985). These *F*-statistics are not exactly distributed in this case, because we do not have identical explanatory variables. The finite sample *F*-statistic is, however, more conservative than either the asymptotic *F*-statistic or the chi-square statistic.[12]

Primary Results

We first estimated equation (3) with one forecast coefficient and a quarterly earnings announcement coefficient, in which all forecasts (regardless of type) were aggregated into a single independent variable. This approach allowed us to assess whether the inclusion of interim forecasts and minimum and maximum projections influences our inferences about the information content of management forecasts.

Table 4 presents results from this estimation. The mean (median) forecast coefficient is 0.24 (0.09). We are able to reject hypothesis H_{01} that the average coefficient equals zero and hypothesis H_{02} that all coefficients jointly equal zero at the 0.01

---

[11] See Schipper and Thompson (1983) and Pownall (1986) for a discussion of the correspondence between using conditional market models and using more traditional residual analysis methods.

[12] The differences across firms in the independent variables are (1) UD_{ikt} takes different values for each disclosure for each firm, and the nonzero values of the variables are on different days, and (2) not all firms have nonzero values for all vectors of management forecast variables used in some estimations. See deJong and Thompson (1989) for a discussion of the finite sample *F*-statistic distribution as a more conservative approximation than either the chi-square or asymptotic *F*-statistic distributions.

Table 4
Tests for Information Content of Management Forecasts

Model: $R_{it}=\alpha_i+\beta_i R_{mt}+\gamma_{i1}MF_{it}+\gamma_{i2}EA_{it}+\epsilon_{it}$.[a]

	MF	EA
F-statistics[b]		
for H_{01}:	20.37***	75.66***
for H_{02}:	3.75***	5.86***
for H_{03}:	MF=EA: 8.66***	
Mean	0.24	0.47
Median	0.09	0.15
Range	−1.68 to 6.18	−0.62 to 5.25
N (%<0)	90 (38%)	90 (19%)

[a] This is the model in equation (3), where R_{it}=return to security i on day t; R_{mt}=return to the value-weighted market index on day t; α_i and β_i=market model parameters; $MF_{it}=UD$ as defined in equation (1) on the day before and the day of a forecast; $EA_{it}=UD$ on the day before and the day of an earnings announcement; γ_{ik}=the stock price response to forecasts ($k=1$) and earnings announcements ($k=2$); and ϵ_{it}=multivariate normal random disturbances.

[b] F-statistics for tests of hypothesis H_{01} have one and 2,143 degrees of freedom. F-statistics for tests of hypothesis H_{02} have I and $(2{,}144-I)$ degrees of freedom, where I is the number of firms in each test.

Hypothesis H_{01}: $\Sigma\gamma_{ik}=0$.
Hypothesis H_{02}: $\gamma_{ik}=0\ \forall i$.
Hypothesis H_{03}: $\bar{\gamma}_{i1}=\bar{\gamma}_{i2}$.

*** Denotes significant at the 0.01 level.

level. These results support previous studies indicating that forecasts are price informative (e.g., Patell 1976; Penman 1980; Waymire 1984). The mean and median earnings announcement coefficients are more positive than the forecast coefficients. The average earnings announcement coefficient is significantly different from zero at the 0.01 level, and significantly differs from the average management forecast coefficient at the 0.05 level.

This last result is not consistent with Pownall and Waymire (1989), who find that management forecasts are more price informative than annual earnings announcements. These differences could be due to difference in samples or sampling periods.

We next disaggregated the forecasts by type (i.e., points, ranges, minimums, and maximums). The model in table 5, panel A includes a coefficient for point forecasts, a coefficient for all other forecasts, and a coefficient for earnings disclosures. We estimated this model to isolate the effects of point forecasts relative to more qualitative forecasts. For all three coefficients, we can reject hypotheses H_{01} and H_{02} at conventional levels. Although the average point forecast coefficient estimate is the most positive of the three, we cannot reject the hypothesis that it equals the average for other forecasts or earnings announcements at 0.10.[13]

---

[13] We also estimated a model that pooled point and range projections into one coefficient and included minimum and maximum forecasts in another. The mean (median) point/range forecast coefficient was 0.42 (0.19), compared with 0.18 (0.06) for the minimum/maximum coefficient. We could not reject hypothesis H_{01} for the points and ranges at the 0.10 level, but we rejected hypothesis H_{01} at the 0.05 level for minimums and maximums. In both cases, we rejected hypothesis H_{02} at the 0.01 level.

Table 5
Tests for Information Content of Management Forecasts Partitioned by Forecast Type

Panel A. Estimating Coefficients for Point Forecasts Separately:

$$\text{Model: } R_{it}=\alpha_i+\beta_i R_{mt}+\gamma_{i1} PT_{it}+\gamma_{i2} OMF_{it}+\gamma_{i3} EA_{it}+\epsilon_{it}.^{a}$$

	Forecast Type		
	Point (PT)	*All Other (OMF)*	*EA*
F-statistics[b]			
for H_{01}:	5.94**	7.51***	75.65***
for H_{02}:	3.55***	2.87***	5.86***
Mean	0.78	0.29	0.47
Median	0.32	0.07	0.14
N	70	89	90

Panel B. Estimating Coefficients for Each Forecast Type Separately:

$$\text{Model: } R_{it}=\alpha_i+\beta_i R_{mt}+\sum_{k=1}^{5}\gamma_{ik} UD_{ikt}+\epsilon_{it}.^{c}$$

	Forecast Type				
	Point	*Range*	*Minimums*	*Maximums*	*EA*
F-statistics					
for H_{01}:	6.20***	3.03*	2.85*	2.12	77.07***
for H_{02}:	3.49***	1.97***	2.00***	2.67***	5.87***
Mean	0.82	0.75	0.23	0.70	0.47
Median	0.46	0.35	0.06	0.10	0.14
N	69	57	87	66	90

[a] This model is a variant of equation (3), where R_{it}, R_{mt}, α_i, β_i, ϵ_{it}, and EA_{it} are defined as in footnote a to table 4; $PT_{it}=UD$ on the day before and the day of a point forecast; $OMF_{it}=UD$ on the day before and the day of a range, minimum, or maximum forecast; and $\gamma_{ik}=$ the stock price response to points ($k=1$), other forms of forecasts ($k=2$), and earnings announcements ($k=3$).

[b] F-statistics for tests of hypothesis H_{01} have one and 2,143 degrees of freedom. F-statistics for tests of hypothesis H_{02} have I and $(2{,}144-I)$ degrees of freedom, where I is the number of firms in each test.

[c] This model is a variant of equation (3), where R_{it}, R_{mt}, α_i, β_i, and ϵ_{it} are defined as in footnote a to table 4; $UD_{ikt}=$ the value defined in equation (1) on the day before and the day of the relevant forecast or earnings announcement and zero otherwise; and γ_{ik} ($k=1$ to 5) is the parameter measuring the firm's stock price response to earnings announcements or to one of the four types of forecasts.

* Denotes significant at the 0.10 level.
** Denotes significant at the 0.05 level.
*** Denotes significant at the 0.01 level.

Panel B of table 5 shows estimation results for a variant of equation (3) with four management forecast coefficients, one each for point, range, minimum, and maximum forecasts. We can reject both hypotheses H_{01} and H_{02} at the 0.01 level for point forecasts. For the minimums, we can reject hypothesis H_{01} weakly (at 0.10) and hypothesis H_{02} strongly (at 0.001). Hypotheses H_{01} and H_{02} can likewise be rejected at identical levels for the range forecasts, and the average coefficient on maximum forecasts is not significantly different from zero at the 0.10 level. No pairwise comparison enables us to reject the null hypothesis of average coefficient equality (H_{03}) at the 0.10 level.

Panel A of table 6 provides estimation results for a model with separate coefficients for interim and annual earnings forecasts, along with a coefficient for earnings announcements. These results strongly support the hypothesis that interim projections are more informative than annual forecasts. We can reject the hypothesis that the mean interim coefficient equals the mean annual coefficient at the 0.001 level. Interim forecasts are significantly more price-informative than earnings announcements, which are significantly more price-informative than annual forecasts, both at the 0.05 level.

Panel B of table 6 reports the estimation results for a model with eight separate forecast coefficients, one for each type of interim forecast and one for each type of annual forecast. This is the most disaggregated data level for which we estimated models. EA results, not reported, are nearly identical to those reported in tables 4, 5, and 6A. The results are rather weak, probably due to small sample sizes. All mean coefficient estimates except interim minimums are positive, but only three of the average coefficients differ significantly from zero. Also, mean coefficient estimates are larger for interim disclosures of three types (points, ranges, and maximums), although the medians are larger for the annual estimates.

The panel B results suggest that the differences between interim and annual management forecasts in panel A are probably driven by the significant difference between interim and annual maximums. Maximums are the only case in panel B in which the interim mean is significantly more positive than the annual mean, and proportionately more maximums are interim forecasts (relative to other forecast types). However, the coefficients underlying the results in panel B are estimated over smaller samples, and therefore are more subject to estimation error than earlier test results. This problem is apparent in the disparity of coefficient estimate means across forecast types. Consequently, these results should be interpreted cautiously.

Our primary results support several conclusions. First, forecasts are significantly related to contemporaneous returns even with the inclusion of more qualitative and interim forecasts not examined in prior studies. Second, forecasts are significantly less informative than earnings announcements, a finding that is not consistent with the results in Pownall and Waymire (1989). Third, no significant differences exist between the price responses to alternative forecast types, although point forecasts are generally associated with estimated coefficients that are more positive and more significant. Fourth, forecasts of interim earnings are significantly more price-informative than annual earnings projections.

III. Diagnostics and Extensions

Price Effects of Minimum and Maximum Forecasts

In table 6, panel B, the sign reversals between the mean and median for annual and interim minimums and interim maximums suggest that the *UD*s for these forecasts may be materially misspecified. To examine the extent of the *UD* misspecification, we assessed the sign of price changes associated with minimum (lower bound) and maximum (upper bound) forecasts. Because these forecasts are truncations of one tail of the earnings distribution, minimum (maximum) forecasts may be associated with

Table 6
Tests for Information Content of Management Forecasts Partitioned by Forecast Period

Panel A. Partitioned by Forecast Period:

Model: $R_{it}=\alpha_i+\beta_i R_{mt}+\gamma_{i1}MFI_{it}+\gamma_{i2}MFA_{it}+\gamma_{i3}EA_{it}+\epsilon_{it}$.[a]

	Forecast Period		
	Interim	*Annual*	*EA*
F-statistics			
for H_{01}:	14.40***	11.11***	75.44***
for H_{02}:	2.88***	2.89***	5.78***
Mean	1.39	0.27	0.47
Median	0.13	0.07	0.14
N	82	87	90

F-statistics for H_{03}:

MFI = MFA F-statistic: 6.69***
MFI = EA F-statistic: 4.37**
MFA = EA F-statistic: 5.75**

Panel B. Partitioned by Both Forecast Period and Forecast Type:

Model: $R_{it}=\alpha_i+\beta_i R_{mt}+\sum_{k=1}^{9}\gamma_{ik}UD_{ikt}+\epsilon_{it}$.[b]

	Forecast Type			
Forecast Period	*Point*	*Range*	*Minimum*	*Maximum*
Interim				
F-statistic for H_{01}	0.36	0.94	5.57**	15.02***
Mean	0.73	1.45	−3.36	6.98
Median	0.17	0.18	0.32	−0.30
N	44	28	59	48
Annual				
F-statistic for H_{01}	0.76	0.52	0.73	21.05***
Mean	0.31	0.76	0.14	1.68
Median	0.37	0.35	−0.002	0.34
N	52	45	82	52

F-statistics for H_{03}:
Annual Minimum = Interim Minimum F-statistic: 5.98**
Annual Maximum = Interim Maximum F-statistic: 8.35***

---

[a] This model is a variant of equation (3), where R_{it} = the return to security i on day t; R_{mt} = the return to the value-weighted market index on day t; α_i and β_i = market model parameters; ϵ_{it} = multivariate normal random disturbances; $MFI_{it}(MFA_{it})=UD$ as defined in equation (1) on the day before and the day of an interim (annual) forecast and zero otherwise; $EA_{it}=UD$ as defined in equation (1) on the day before and the day of an earnings announcement and zero otherwise; γ_{ik} = the stock price response to interim ($k=1$) and annual ($k=2$) forecasts and to earnings announcements ($k=3$).

[b] This model is a variant of equation (3), where R_{it}, R_{mt}, α_i, β_i, and ϵ_{it} are defined as in footnote a to this table; UD_{ikt} ($k=1$ to 9) = UD as defined in equation (1) on the day before and the day of the relevant interim or annual forecast or quarterly earnings announcement date and zero otherwise; γ_{ik} ($k=1$ to 5) = the firm's stock price response to quarterly earnings announcements and each type of forecast.

** Denotes significant at the 0.05 level.
*** Denotes significant at the 0.01 level.

systematically positive (negative) abnormal returns. Baginski et al. (1992) report evidence consistent with this conjecture for a sample of post year-end estimates.

Accordingly, we estimated a model in which the disclosures were partitioned by type, and 0 – 1 dummy variables were used for minimums and maximums. In this specification, the estimated coefficient measures the mean abnormal return on days when the variable assumes a value of one. We expected the estimated coefficient to be positive (negative) for minimum (maximum) forecasts. Panel A of table 7 shows results of this specification for the full sample, and panel B provides results using only minimums (maximums) greater (less) than or equal to the prior analyst forecast. EA results, not reported, are nearly idenfical to those reported in tables 4, 5, and 6A.

The results in panel A indicate that the maximums are associated with highly significant negative mean abnormal returns. The mean coefficient estimate for the minimum forecasts is positive but not significant. Consistent with Baginski et al. (1992), the price effects are more pronounced for maximum forecasts than for minimum forecasts. Results in panel B are generally consistent with those in panel A, except that minimum forecasts on average are price-informative when minimums that are less than the analyst forecast are excluded.

Initial Forecasts vs. Management Forecast Revisions

Of the 1,252 sample disclosures, 844 are initial forecasts and the remaining 408 are revisions of prior forecasts. Also, 355 of the forecast revisions pertain to annual projections, and only 53 apply to previous interim projections. Additional tests suggested that our results on point projections were strongly influenced by disclosure timing. We estimated a model with forecasts partitioned by type using only the initial forecasts and found that the mean coefficient estimate for point projections declined from 0.82 to 0.53. Further, the mean point coefficient estimate for only initial forecasts was not significant at the 0.10 level. In contrast, the mean point coefficient estimate for only final forecasts was 1.10, which is significant at the 0.01 level. Since there are 161 point forecasts in the initial forecast sample and 189 in the final forecast sample, this result is probably not the artifact of gross disparities in sample size.

Robustness to Alternative Methodologies

We performed two additional tests to evaluate the robustness of our results to alternative methods for estimating information effects. Using a diagonal covariance matrix in the SUR tests supports the same inferencecs, Also, using cross-section regressions of two-day abnormal returns on *UD* for each disclosure type did not alter the results.

Other Issues

We conducted several additional tests to assess the impact of various data partitions or other methodological changes. First, we identified 37 management forecasts that were released simultaneously with dividend changes, stock splits, or announcements about mergers and acquisitions. Omitting these observations had no material effect on our conclusions. Second, several annual forecast observations released in the fourth fiscal quarter were disclosed with fourth-quarter forecasts. Deleting either the fourth-quarter or the annual forecasts in these cases had little effect on our conclusions.

Finally, the *DJNRS* followed a more restrictive editorial policy from its inception in

Table 7
Diagnostic Analyses of the Information Content of Minimum and Maximum Forecasts (Using Dummy Variables)

Model: $R_{it}=\alpha_i+\beta_i R_{mt}+\gamma_{i1} PT_{it}+\gamma_{i2} RG_{it}+\gamma_{i3} DVMIN_{it}+\gamma_{i4} DVMAX+\gamma_{i5} EA_{it}+\epsilon_{it}$.[a]

Panel A. Estimated on Full Data Set:[b]

	PT	*RG*	*DVMIN*	*DVMAX*
F-statistic for H_{01}[c]	4.67**	1.79	0.17	51.07***
Mean	0.72	0.58	0.0004	−0.009
Median	0.29	0.09	0.002	−0.008
Range	−23.10 to 21.99	−13.26 to 16.84	−0.05 to 0.05	−0.05 to 0.04
N	69	57	87	66
N (%) positive & significant at 0.1 level			14 (16.1)	2 (3.0)
N (%) negative & significant at 0.1 level			9 (10.3)	22 (33.3)

Panel B. Estimated Using Only Unambiguous Observations:[d]

	PT	*RG*	*DVMIN*	*DVMAX*
F-statistic for H_{01}	3.52*	2.49	7.66***	14.26***
Mean	0.68	0.70	0.0049	−0.015
Median	0.23	0.31	0.003	−0.01
Range	−23.25 to 22.11	−13.25 to 16.88	−0.04 to 0.04	−0.30 to 0.04
N	68	55	52	56
N (%) positive & significant at 0.1 level			12 (23.1)	1 (1.8)
N (%) negative & significant at 0.1 level			3 (5.8)	17 (30.4)

[a] This model is a variant of equation (3), where R_{it}, R_{mt}, α_i, β_i, ϵ_{it}, PT_{it}, and EA_{it} are defined as in footnote a to table 4; $RG_{it}=UD$ on the day before and the day of a range forecast; $DVMIN=1$ on the day before and the day of a minimum forecast; $DVMAX=1$ on the day before and the day of a maximum forecast; and γ_{ik} = the stock price response to the various forecast types specified in this way.

[b] Results on this specification of the minimum and maximum vectors do not differ materially when estimated on interim and annual observations separately.

[c] *F*-statistics for tests of hypothesis H_{01} have one and 2,143 degrees of freedom. *F*-tests of the null hypothesis that the firm-specific dummy variable coefficients for minimums and maximums are jointly equal to zero reject the null at the 0.01 level. In addition, *F*-tests of the null hypotheses that the *DVMIN* and *DVMAX* coefficients are equal to each other on average across the sample or in within-firm comparisons reject the null at the 0.01 level.

[d] Minimum forecasts should be unambiguously good news when the benchmark is above the current analyst forecast. Similarly maximum forecasts should be unambiguously bad news when the benchmark is below the current analyst forecast. Therefore we define unambiguous observations as minimum forecasts greater than or equal to the current analyst forecast and maximum forecasts less than or equal to the current analyst forecast.

* Denotes significant at the 0.10 level.
** Denotes significant at the 0.05 level.
*** Denotes significant at the 0.01 level.

July 1979 until December 1983, and a more liberal editorial policy from January 19
until the present. We therefore have a potentially more restricted sample of disclosur
before 1984. Omitting disclosures between July 1, 1979 and December 31, 1983 also h
no material effect on our conclusions, although smaller sample sizes resulted in weak
inferences in the disaggregated analyses.

IV. Summary

In this paper we examine the stock price effects of management earnings forecas
that differ by form and horizon. Prior studies have provided only limited evidence
the information content of alternative forecast disclosures. Our tests yield several fin
ings. First, forecasts have information content, in general, even when includi
forecast forms and horizons not analyzed in prior studies. Second, for our full sampl
forecasts are less informative than earnings announcements. Third, differences b
tween point forecasts and other forecasts are generally not significant at convention
levels. Fourth, interim earnings forecasts are significantly more informative th
annual forecasts.

Our results suggest several issues for further research. First, broader (and large
management forecast samples may yield additional insights on the information co
tent of management forecasts. For example, the impact of range "spreads" on inf
mativeness could be evaluated, as could the information content of manageme
forecast revisions. Also, are managers more likely to disclose revisions when they ha
more negative information than that implied by an outstanding forecast? Second, whi
an extensive literature exists on the accuracy of management forecasts (see King et
1990), additional study could be focused on the accuracy of alternative forecast type
In particular, do analysts update their projections around the release of alternati
types of management forecasts? Analyst forecast revisions and management foreca
accuracy could also be linked to stock price formation as in Jennings (1987) a
McNichols (1989).

Third, more direct tests of managers' incentives to disclose qualitative forecas
would be useful. For example, do firms that issue relatively more qualitative projectio
have less stable earnings series or other distinguishing characteristics? Fourth, our tes
could be extended to other disclosures where the form of disclosure may be central
its informativeness, such as the disclosure and valuation of contingent liabilities,
issue of current interest to the FASB.

References

Ajinkya, B., and M. Gift. 1984. Corporate managers' earnings forecasts and symmetrical adjustme
of market expectations. *Journal of Accounting Research* 22 (Autumn): 425–44.

Baginski, S., E. Conrad, and J. Hassell. 1993. The effects of management forecast precision on equi
pricing and on the assessment of earnings uncertainty. *The Accounting Review* 68 (Octobe
913–27.

———, J. Hassell, and G. Waymire. 1992. The nature of news communicated in preliminary earnin
estimates. Working paper, Emory University.

deJong, P., and R. Thompson. 1989. Testing linear hypotheses in the SUR framework with identic
explanatory variables. *Research in Finance* monograph series 9. Greenwich, CT: JAI Press.

Hoskins, R., J. Hughes, and W. Ricks. 1986. Evidence on the incremental information content
additional firm disclosures made concurrently with earnings. *Journal of Accounting Resear*
24 (Supplement): 1–32.

Jennings, R. 1987. Unsystematic security price movements, management earnings forecasts, and revisions in consensus analyst earnings forecasts. *Journal of Accounting Research* 25 (Spring): 90–110.

King, R., G. Pownall, and G. Waymire. 1990. Expectations adjustment via timely earnings forecast disclosure: Review, synthesis, and suggestions for future research. *Journal of Accounting Literature* 9: 113–44.

Lev, B., and S. Penman. 1990. Voluntary forecast disclosure, non-disclosure, and stock prices. *Journal of Accounting Research* 28 (Spring): 49–76.

McNichols, M. 1989. Evidence of informational asymmetries from management earnings forecasts and stock returns. *The Accounting Review* 64 (January): 1–27.

Patell, J. 1976. Corporate forecasts of earnings per share and stock price behavior: Empirical tests. *Journal of Accounting Research* 14 (Autumn): 246–76.

Penman, S. 1980. An empirical investigation of the voluntary disclosure of corporate earnings forecasts. *Journal of Accounting Research* 18 (Spring): 132–60.

Philbrick, D., and W. Ricks. 1991. Using Value Line and IBES analyst forecasts in accounting research. *Journal of Accounting Research* 29 (Autumn): 397–417.

Pownall, G. 1986. An empirical analysis of the regulation of the defense contracting industry: The Cost Accounting Standards Board. *Journal of Accounting Research* 24 (Autumn): 291–315.

———, and G. Waymire. 1989. Voluntary disclosure credibility and securities prices: Evidence from management earnings forecasts, 1969–1973. *Journal of Accounting Research* 27 (Autumn): 227–45.

Schipper, K., and R. Thompson. 1983. The impact of merger-related regulations on the shareholders of acquiring firms. *Journal of Accounting Research* 21 (Spring): 184–221.

———, and ———. 1985. The impact of merger-related regulations using exact distributions of test statistics. *Journal of Accounting Research* 23 (Spring): 408–15.

Thompson, R., C. Olsen, and J. Dietrich. 1987. Attributes of news about firms: An analysis of firm-specific news reported in *The Wall Street Journal Index*. *Journal of Accounting Research* 25 (Autumn): 245–74.

Waymire, G. 1984. Additional evidence on the information content of management earnings forecasts. *Journal of Accounting Research* 22 (Autumn): 703–18.

Zellner, A. 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* 57 (June): 348–68.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 913–927

The Effects of Management Forecast Precision on Equity Pricing and on the Assessment of Earnings Uncertainty

Stephen P. Baginski
Florida State University
Edward J. Conrad
The University of Akron
John M. Hassell
The University of Texas at Arlington

SYNOPSIS AND INTRODUCTION: This study examines the effects of management forecast precision (i.e., lack of uncertainty) on equity pricing and the assessment of earnings uncertainty. Kim and Verrecchia (1991) modeled the price reaction to the public release of information as a positive function of both the unexpected component of the information and the information's precision. We test these predictions with a sample of 868 management forecasts for 1983–1986 annual and interim earnings. The use of management forecasts rather than actual earnings to test the precision hypothesis has the distinct advantage that the level of forecast precision is not directly regulated and thus may vary across forecasts. Further, managers explicitly disclose their level of uncertainty. Both this study and Pownall et al. (1993) document that most forecasts are open-interval (minimums and maximums), closed-interval (ranges), or general impressions rather than point estimates.

The method used to test the precision hypothesis removes restrictions on the traditional regression of unexpected returns on unexpected earnings. Specifically, the slope and intercept coefficients that map unexpected earnings into unexpected returns can vary in the cross-section as a function

We thank Greg Waymire, Grace Pownall, Wanda Wallace, Jeff Clark, Tom Schaefer, Ken Lorek, Al Bathke, John Neill, two anonymous referees, and workshop participants at the University of Texas at Arlington, Texas A&M University, St. Louis University, and Florida State University for comments on earlier drafts of this paper. Also, we thank IBES, Inc. for the analyst forecast data.

Submitted October 1991.
Accepted March 1993.

of forecast precision. Our results support a direct relation between forecast precision and the importance of management forecasts for security pricing.

Holthausen and Verrecchia (1990) and Morse et al. (1991) modeled a decrease in investors' consensus as a positive function of the magnitude of signal surprise and the dispersion of the perceived precision of the signal. We examine these predictions with a sample of 221 point and closed-interval (range) forecasts. We calculate whether the range of outcomes disclosed by a manager exceeds the range of Institutional Brokers Estimate System (IBES) analyst forecasts. We find this variable and the magnitude of unexpected security returns (a proxy for signal surprise) to be positively associated with increases in the standard deviation of IBES analyst forecasts. Morse et al. (1991) found the hypothesized relation between signal surprise and increase in analyst forecast variance, but were unable to separate the precision effect from the signal surprise effect. Managers' explicit labeling of forecasts as more uncertain through range disclosure permits the direct calculation of management forecast precision relative to analyst forecast precision.

Our tests involve joint hypotheses of the effects of forecast precision on security prices and the credibility of managers' disclosures of forecast precision. Ajinkya and Gift (1984) developed and tested the "expectations adjustment hypothesis" which posits sufficient incentives for credible, symmetric forecast disclosure. King et al. (1990) argued that expectations adjustment also suggests credible labeling of the precision of forecasts.

Key Words: *Forecast precision, Belief dispersion, Expectations adjustment.*

Data Availability: *Additional papers using portions of these data are in progress. Requests for data available from public sources used in this study should indicate the intended use.*

THE remainder of this paper is organized as follows. Section I motivates the hypotheses to be tested. Section II presents sample selection, data, and research design. Section III discusses results, and section IV summarizes the article.

I. Hypotheses

Expectations Adjustment

Ajinkya and Gift (1984) posit a symmetric relation between management forecast news and security prices using the expectations adjustment hypothesis. Under this hypothesis, managers are motivated to disclose forecasts voluntarily when they wish to align investor expectations with their own. As King et al. (1990) argue, the driving force behind this motivation is the managers' desire to reduce transaction costs arising from private information acquisition and resulting information asymmetry. Diamond (1985) provides a formal analysis of how public disclosure reduces the incentives to acquire private information, the ability to realize trading gains, and the resulting transaction costs. Ajinkya and Gift (1984), Waymire (1984) and McNichols (1989) find evidence consistent with the expectations adjustment hypothesis for point and range predictions of earnings and earnings per share (EPS).

More recently, Pownall et al. (1993) document that the majority of management forecasts are not point projections of earnings or earnings per share, and convey various levels of precision as indicated by ranges or conditioning phrases.[1] King et al. (1990) argue that managers will credibly convey the precision of their estimates to maintain a reputation for high quality disclosure. If expectations adjustment is descriptive of management behavior and managers explicitly label forecast precision, then recent models provide the basis for empirical predictions of the effects of forecast precision on security prices and changes in belief consensus.[2]

Precision and Security Pricing

Kim and Verrecchia (1991) examine a two-period rational expectations model where traders are assumed to be diversely informed and differ in the precision of their private prior information. They find that the price reaction to the unexpected portion of a disclosure is an increasing function of its relative importance across the posterior beliefs of traders. The relative importance is positively related to the precision of the announcement and inversely related to the precision of preannouncement information. This theoretical result is consistent with intuition and the empirical findings in Atiase (1985).

Since the expectations adjustment scenario attributes credible labeling behavior to managers, Kim and Verrecchia's study implies that price informativeness for the unexpected component of a forecast declines with the degree of forecast imprecision. The directional hypothesis (stated in the alternative form) is:

> H1: The strength of the association between the signal contained in a management forecast of earnings and stock price response to that signal is negatively related to the forecast's imprecision.

Precision and Changes in Belief Consensus

Morse et al. (1991) present a Bayesian belief revision model that links the surprise content of a signal and its precision to changes in belief consensus. The predictions of this model are that both surprise content and dispersion of perceived signal precision are positively related to an increase in the dispersion of beliefs. Holthausen and Verrecchia (1990) also model reductions in consensus as a positive function of disagreement over a given signal's implications.

If managers credibly disclose the uncertainty present in their forecasts, and their uncertainty is greater than prevailing uncertainty among traders, then disagreement about expected earnings should increase. Our directional hypotheses based on this prior research (stated in the alternative form) are:

---

[1] Examples of phrases used by managers to condition their forecasts include "less than," "greater than," "at least," "no more than," etc.

[2] We do not intend to imply that expectations adjustment is the only scenario that leads to credible forecasting or that our empirical results are consistent only with the expectations adjustment motivation for forecasting. We chose expectations adjustment because of the accumulated empirical evidence in its favor and because King et al. (1990) conveniently extend expectations adjustment to a prediction about credible labeling of forecast precision, the object of our empirical analysis.

H2: Other things being equal, the release of managers' forecasts that indicate a greater uncertainty about future earnings than that indicated by prevailing analysts' forecasts of earnings leads to an increase in the dispersion of analysts' forecasts.

H3: Other things being equal, the release of managers' forecasts that indicate a greater surprise about future earnings leads to an increase in the dispersion of analysts' forecasts.

We are primarily interested in testing hypothesis H2 because it links precision with changes in beliefs. We test hypothesis H3 as a means of extending Morse et al.'s (1991) work, which rejects the hypothesis H3 null for actual earnings releases.

II. Sample Selection and Research Design

Sample Selection

Management forecasts of quarterly and annual earnings and earnings per share were obtained from the *Dow Jones News Retrieval Service* (*DJNRS*) for the period from January 1, 1983 to December 31, 1986. The *DJNRS* contains articles published in *The Wall Street Journal* and *Barrons*, as well as unpublished announcements appearing on the Broad Tape. This data source was queried using key words suggesting that the news item was a forecast by a corporate official.[3]

Table 1 summarizes the steps employed in the data filtering process and the number of forecasts and firms remaining in the sample. The final sample consists of 868 forecasts issued by 569 firms. Forecasts of annual results total 696 (80.2 percent), while forecasts of interim results total 172 (19.8 percent).[4] The management forecast data include point, range (closed-interval), and open-interval estimates of EPS. Point forecasts do not contain conditioning phrases such as "greater than," "less than," "no more than," or "at least." Open-interval forecasts include one of these types of phrases so that the interval of potential outcomes indicated by the forecast is unbounded at one end. Range forecasts (closed-interval) have both ends of the interval bounded, usually through language such as "between \$x and \$y." Prior management forecast studies have examined primarily point and range estimates of earnings (converting range forecasts to the midpoint of the range).[5]

---

[3] This approach is analogous to that used by Hassell and Jennings (1986) and McNichols (1989). The data base search is conducted using key words that do not discriminate *a priori* based on factors, including corporate characteristics and forecast content, other than the key words contained in the article's headline.

[4] Examination of those firms excluded by the COMPUTSTAT and CRSP data criteria reveals that they are almost exclusively firms traded over the counter. IBES quarterly forecasts began in October 1983 and are not available for as many firms as are annual forecasts.

[5] Examples of these studies include Penman (1980), Waymire (1984), Hassell and Jennings (1986), Jennings (1987), and Baginski and Hassell (1990). Pownall et al. (1993) use a much broader set of forecasts for a sample of repeat forecasters that include open-interval forecasts. Also, Patell (1976) and Cox (1985) make use of the information available in open-interval estimates of the "greater than" or "less than" form. Neither of these studies examines the hypotheses in this paper, although Patell (1976) does disclose mean cross-sectional unexpected returns at forecast dates for an open-interval sample.

Table 1
Sample Selection and Filtering Process
(1983–1986 Sampling Period)

Filtering Process	*Forecasts*
Observations identified by *DJNRS* search	5,031
Foreign firm forecasts; forecasts not made by a corporate official; forecasts not firm specific; forecasts for period other than a quarter or annual period	(1,233)
Insufficient COMPUSTAT and CRSP data (includes non-listing)	(1,648)
General impression forecasts (unexpected earnings not determinable)	(222)
Issued before beginning or after end of quarterly or annual period	(431)
Not followed by IBES financial analysts	(629)
Final sample	868

Distribution by Year	*Firms*	*Forecasts*
1983	138	138
1984	188	197
1985	224	255
1986	222	278
1983–1986	569	868

Distribution by Forecast Horizon	*Forecasts*	*Percent*
Annual	696	80.2%
Interim (quarters 1–3)	172	19.8%
Total	868	100.0%

Research Design for Hypothesis H1

To examine the effects of precision on the mapping of unexpected earnings into security prices, the following model was estimated cross-sectionally using ordinary least squares regression:

$$CPE_i = b_0 + b_1(UE_i) + b_2(IMPRECISE_i) + b_3(IMPRECISE_i{*}UE_i) + e_i. \qquad (1)$$

Equation (1) is a variant of the model traditionally used to relate unexpected earnings (*UE*) to cumulative market model prediction errors (*CPE*s). The intercept and slope coefficients are allowed to vary in the cross-section as a function of management forecast imprecision (*IMPRECISE*). *IMPRECISE* equals zero for point forecasts, one for closed-interval forecasts, and two for open-interval forecasts. These values correspond to the levels of uncertainty conveyed in the forecast as measured by the potential outcomes indicated by the manager.[6]

CPE measures the security price reaction to the management forecast over a three-day event day interval centered on the *DJNRS* date (day 0) using standard market model estimation techniques:

$$CPE_i = \sum_{t=-1}^{1} PE_{it}; \tag{2}$$

where $PE_{it} = [R_{it} - (\hat{\alpha}_i + \hat{\beta}_i R_{mt})]$; R_{it} = return for firm i on day t; R_{mt} = return on an equally-weighted market portfolio on day t; and $\hat{\alpha}_i$ and $\hat{\beta}_i$ = ordinary least squares intercept and slope coefficients of the familiar market model estimated over days −230 to −31 in event time.

Unexpected earnings (*UE*) is measured by comparing the management forecast to the preceding analyst forecast and scaling by price:

$$UE_i = (MF_i - AF_i)/Price_i, \tag{3}$$

where MF_i = the management forecast of EPS for firm i; AF_i = the median analyst forecast of EPS for firm i obtained from IBES for the month preceding the management forecasts; and $Price_i$ = day −2 security price for firm i.

For point estimates of EPS, *MF* is provided explicitly by the manager. For closed-interval (range) forecasts, we averaged the two endpoints provided by the manager to obtain *MF*. For minimum and maximum (open-interval) forecasts, we used the disclosed upper or lower bound as our estimate of *MF*.

Our measurement scheme for *MF* (and hence, *UE*) is a function of the level of forecast imprecision. In the appendix, we replicate our tests using only those cases where we can obtain an unambiguous measure of *UE*. The measurement scheme used in the appendix does not depend on the level of management forecast precision. Our results reported in section III are robust with respect to these alternative tests. It also appears that the *MF* measurement scheme used in equation (3) for closed-interval and open-interval management forecasts is reasonable. Of particular interest is the finding that, when confronted with a closed-interval management forecast, investors seem to use the middle of the closed-interval as the point estimate.

The well-documented information content of earnings hypothesis predicts $b_1 > 0$. The alternative to hypothesis H1 predicts $b_3 < 0$ to indicate that imprecise forecasts are associated with a weaker relation between unexpected earnings and security prices. In equation (1), a negative b_3 multiplied by *UE* shifts the price response to the forecast in a direction opposite the sign of *UE*. We do not make specific predictions about the intercept, b_0, and the intercept shift, b_2, coefficients.[7]

---

[6] Point forecasts convey only one potential outcome. Closed-interval forecasts convey several possible outcomes, but the possibilities are bounded at both ends. Open-interval forecasts convey an infinite number of outcomes to one side of the distribution.

[7] A significant intercept has often been interpreted as indicative of the pricing effects of the act of disclosure. We have no expectation for the effects of imprecision *per se* on pricing. A b_2 coefficient that was of the opposite sign of b_0 would, however, be consistent with the notion that any pricing effects are dampened in the case of imprecise forecasts.

Research Design for Hypotheses H2 and H3

Alternative hypotheses H2 and H3 predict a positive relation between increases in the dispersion of investor beliefs and both (1) the imprecision of a management forecast relative to investor beliefs and (2) the surprise content of the management forecast. IBES discloses the standard deviation across analysts' forecasts of EPS at any given time. We adopt the standard deviation as our proxy for the dispersion of investor beliefs.[8]

We chose the month preceding the management forecast release as the first measurement point for the standard deviation of analysts' EPS forecasts ($AFSTD_{PRE}$) and the month immediately following release as the second measurement point ($AFSTD_{POST}$). By the month after release, analysts should have revised their expectations for earnings.[9] Further, we eliminated those observations for which $AFSTD_{POST}$ did not clearly precede the actual earnings release to ensure that $AFSTD_{PRE}$ and $AFSTD_{POST}$ relate to the same forecasting period.

For certain tests, we simply noted whether the standard deviation of analyst forecasts increased or decreased. For other tests we computed the percentage decrease beteen the two dates ($SDDROP$):[10]

$$SDDROP_i = (AFSTD_{PRE,i} - AFSTD_{POST,i}) / AFSTD_{PRE,i}, \tag{4}$$

where $AFSTD_{PRE,i}$ = the standard deviation in IBES forecasts in the month preceding the month of management forecast release for firm i, and $AFSTD_{POST,i}$ = the standard deviation in IBES forecasts in the month following the month of management forecast release for firm i.

Brown et al. (1985) document a decrease in the standard deviation of IBES forecasts as the actual earnings date approaches. Thus, we expect standard deviation decreases to outnumber increases and the mean of $SDDROP$ to be positive.

The two variables hypothesized as partially explaining a change in the IBES standard deviation are (1) the imprecision of a management forecast relative to investor beliefs, and (2) the surprise content of the management forecast. The former is difficult to proxy because managers do not release the standard deviation that characterizes their distribution. The width of the range of beliefs disclosed in managers' point and closed-interval forecasts is the only measure of dispersion explicitly provided by management that can be matched to a dispersion measure on the IBES tape. A similar width measure can be obtained for analysts since the IBES tape provides the lowest and highest analyst forecast. Thus, we limited our tests of hypotheses H2 and H3 to point and closed-interval management forecasts.[11] We computed the following two variables:

---

[8] Morse et al. (1991) use the square of this measure as their proxy.

[9] Empirical evidence suggests that the analyst revision process subsequent to an earnings-related announcement takes about one to two months. Jennings (1987) and Baginski and Hassell (1990) document analyst revision activity immediately after management forecast release. Morse et al. (1991) provide evidence consistent with analysts revising over a two month period subsequent to an actual earnings release. Baginski and Hassell (1990) improve the power of a model explaining analyst revisions by extending the revision period beyond one month. Hassell and Jennings (1986) report that, on average, it takes nine weeks for analyst forecasts to become more accurate than prior management forecasts.

[10] Ajinkya et al. (1991) deflate the standard deviation by the mean analyst forecast to create the scale-free coefficient of variation for use in cross-sectional regression. Since our measure is a percentage change in standard deviation, it is scale-free by construction. Nonetheless, we replicated our analysis by deflating all standard deviations by the mean forecast before computing $SDDROP$. The conclusions we report in this paper are not different from those obtained using this alternative $SDDROP$ definition.

$$MFWIDE_i = MFHI_i - MFLO_i, \tag{5}$$

$$AFWIDE_i = AFHI_i - AFLO_i, \tag{6}$$

where $MFHI_i$ = the high point of the management range forecast for firm i; $MFLO_i$ = the low point of the management range forecast for firm i; $AFHI_i$ = the IBES high analyst forecast for firm i; and $AFLO_i$ = the IBES low analyst forecast for firm i. In the case of point management forecasts, *MFHI* equals *MFLO*, and *MFWIDE* equals zero.

For certain tests, we note whether *MFWIDE* is greater or less than *AFWIDE*. For other tests, we compute a continuous measure of the relative width of the management forecast:

$$\begin{aligned} &RELWIDE_i = MFWIDE_i / AFWIDE_i, && \text{for } AFWIDE \text{ not equal to zero;} \\ &RELWIDE_i = 1, && \text{for } AFWIDE = MFWIDE = 0. \end{aligned} \tag{7}$$

If *AFWIDE* and *MFWIDE* are not equal and *AFWIDE* equals zero, the observation is discarded.[12]

Finally, the surprise content of a management forecast was proxied by standardizing and squaring the unexpected market return. CPE_i from equation (2) was squared and standardized by dividing by the variance of the market model residual during the market model estimation period (event days −230 through −31):[13]

$$SQCPE_i = CPE_i^2 / \text{variance of market model estimation residual.} \tag{8}$$

Our primary test of hypotheses H2 and H3 involves an ordinary least squares estimation of the following cross-sectional model:

$$SDDROP_i = a_0 + a_1(SQCPE_i) + a_2(RELWIDE_i) + e_i. \tag{9}$$

The alternative to hypothesis H2 predicts that $a_2 < 0$, indicating that when the management forecast range is wide relative to the analyst range, the normal decrease in analyst forecast dispersion is attenuated. The alternative to hypothesis H3 predicts that $a_1 < 0$, indicating that larger surprise forecasts also attenuate the decrease in analyst dispersion. Given the results of prior research, we expect the intercept to be positive ($a_0 > 0$), indicating a decrease in analyst forecast dispersion after controlling for both explanatory variables.[14]

---

[11] Even though not disclosed by management (and hence, subject to great uncertainty), an implied upper (lower) bound to minimum (maximum) forecasts is present. There is no way of determining whether the implied width of the forecast is greater than or less than the range of IBES forecasts. Therefore, we limit our analysis to forecasts with no open ends.

[12] Only two observations are lost due to *RELWIDE*'s being undefined. Since these two observations are instances where the management forecast is wider than the analysts' forecasts, we repeated the tests with these two cases arbitrarily set to the largest value of *RELWIDE*. The results were not different than what we report. Also, since the *RELWIDE* proxy has not been used in prior research, we examined the effects of various truncation schemes on the results of the regression. Since conclusions do not change for the various truncation schemes, we report results in table 3 using no truncation for *RELWIDE*.

[13] This procedure is similar in spirit to the method used in Atiase (1985) to examine the relation between surprise and the predisclosure information environment, and in Beaver (1968) to measure surprise endogenously. We used the standardization method outlined in Patell (1976) with the exception of the minor adjustment for variance in the market return outside of the estimation period. This variable is large in some instances due to a small denominator. Our results in tests described later are unaffected by various truncation schemes.

[14] Ziebart (1990) finds no evidence of such a decline at the individual firm level.

Table 2
Tests of the Effect of Forecast Imprecision on the Relation Between Unexpected Earnings and Unexpected Returns

Panel A. Descriptive Statistics:

Continuous Variables	*Mean*	*Median*	*Standard Deviation*
Cumulative prediction error (*CPE*):			
Annual ($n = 696$)	−0.0104	−0.0040	0.0578
Interim ($n = 172$)	−0.0196	−0.0051	0.0782
Unexpected earnings (*UE*):			
Annual ($n = 696$)	0.0013	−0.0008	0.0602
Interim ($n = 172$)	0.0015	−0.0014	0.0277

Classification Variable	*Annual Forecasts*	(%)	*Interim Forecasts*	(%)
Point, *IMPRECISE* = 0	156	(22.4)	40	(23.3)
Closed-interval, *IMPRECISE* = 1	166	(23.9)	24	(13.9)
Open-interval, *IMPRECISE* = 2	374	(53.7)	108	(62.8)
Total	696	(100.0)	172	(100.0)

Panel B. Regression Tests:

Model: $CPE_i = b_0 + b_1(UE_i) + b_2(IMPRECISE_i) + b_3(IMPRECISE_i{*}UE_i) + e_i$.

	Annual (n = 696)		*Interim* (n = 172)	
Coefficient	*Estimate*	*t-statistic*	*Estimate*	*t-statistic*
b_0	−0.014	−3.39**	−0.016	−1.41
b_1	0.255	3.90**	2.360	2.70**
b_2	0.002	0.90	−0.002	−0.31
b_3	−0.095	−2.22*	−0.999	−2.12*
R^2	0.0268		0.0579	
Adjusted R^2	0.0226		0.0411	

* Significant at the 0.05 level in a one-tailed test.
** Significant at the 0.01 level in a one-tailed test.

III. Results

Tests of Hypothesis H1 (The Effects of Precision on Forecast Pricing)

Table 2 presents the results of tests of hypothesis H1. All results are presented for 696 annual forecasts and 172 interim forecasts (quarters one through three) separately. Median cumulative prediction errors and median unexpected earnings are negative for both annual and interim forecasts. The "good news" management forecast tendency documented in Penman (1980) and Patell (1976) does not appear to exist for this sample. In contrast to most prior research, our sample is dominated by open-interval (minimum and maximum) forecasts (53.7 percent of annual forecasts and 62.8 percent of interim forecasts). Point forecasts comprise less than a quarter of the sample.

Consistent with prior management forecast research, coefficient b_1 is significantly positive, indicating the expected relation between unexpected earnings derived from management forecasts and unexpected returns. Coefficient b_3 on the forecast precision by unexpected earnings interaction is negative as predicted by the H1 alternative hypothesis and is significantly different from zero.[15] This result indicates that the strength of the mapping of unexpected earnings into unexpected returns decreases when forecasts become less precise.[16]

Tests of Hypotheses H2 and H3 (The Relation of Relative Management Uncertainty and Forecast Surprise to Changes in Dispersion of Beliefs)

Table 3 presents the results of tests of hypotheses H2 and H3. There are 234 point and closed-interval forecasts with complete analyst forecast standard deviation data. Zero denominators for *SDDROP* or *RELWIDE* require a further reduction of the sample to 214 forecasts.[17] Due to a small number of interim observations, and because we have no reason to believe that hypotheses H2 and H3 are dependent on whether the population is annual or interim management forecasts, we include only annual forecasts in the analysis.

The results of the regression in panel A of table 3 are as expected. The intercept term a_0 is positive and significant, indicating a tendency for analyst forecast standard deviation to decrease through time after controlling for the independent variables.[18] Coefficient a_2 is negative and significant. This result is consistent with alternative hypothesis H2's prediction that when a manager releases a forecast with a relatively wide range, analysts' beliefs become more dispersed. Coefficient a_1 is negative and significant; a result consistent with alternative hypothesis H3 that when there is a greater surprise, analysts' beliefs become more dispersed.[19]

The results reported in panels B and C reexamine the univariate relation between relative management forecast width and change in analyst forecast standard deviation

---

[15] We used collinearity diagnostics provided by Belsley et al. (1980) for all regressions in this study. In no case was a linear dependency detected.

[16] These results were replicated for a larger sample that employed a random walk expectation when an analyst forecast was not available. For annual forecasts, observations totalled 981, the *t*-statistic on b_3 equalled -5.18 ($p<0.01$), and R^2 was 0.0463. For interim forecasts, observations totalled 516, the *t*-statistic on b_3 equalled -2.72 ($p<0.01$), and R^2 was 0.0551. We also used two subsamples of (1) only open-interval forecasts and (2) only closed-interval forecasts to examine whether a further refinement of our *IMPRECISE* measure could be achieved by defining management precision relative to financial analysts' precision. There is a slight improvement in the already significant regression statistics if closed-interval forecasts (*IMPRECISE*=1) are further partitioned on whether the closed-interval is wider than the interval defined by the IBES high and low analyst forecast.

[17] A zero denominator for *RELWIDE* does not eliminate an observation if the numerator is also zero. In that case, *RELWIDE* is equal to one (see eq. [7]).

[18] The mean (median) *SDDROP* is 0.0027 (0.0769) for the sample.

[19] The observations in this analysis differ in the number of analysts following the firm. Requiring a similar number of analysts for all observations would substantially reduce the sample size and hence, the power of the test. To examine the effects of analyst following on coefficients a_0, a_1, and a_2, we reestimated the regression with intercept shift and slope shift coefficients on the number of analysts, and number of analysts by SQCPE interaction, and the number of analysts by *RELWIDE* interaction. This approach uses all data in the regression and preserves the power of the test. Not one of the intercept or slope shift coefficients was significantly different from zero in this unrestricted regression, indicating that a_0, a_1, and a_2 are not functions of analyst following.

Table 3

Tests of the Effects of Relative Forecast Imprecision and Forecast Surprise on Changes in the Dispersion of Investors' Beliefs

Panel A. Regression Test (n = 214):

Model: $SDDROP_t =$	a_0	$+\ a_1(SQCPE_t)$	$+\ a_2(RELWIDE_t)$	$+\ e_t$.
Estimate	0.132	−0.028	−0.176	
t-statistic	3.30**	−2.52**	−5.66**	
$R^2 = 0.1496$				
Adjusted $R^2 = 0.1415$				

Panel B. Test of Differences in SDDROP for Wider vs. Narrower Manager Ranges:

Relation of Management to Analyst Range	*Mean SDDROP*
Management Range Wider ($n = 24$)	−0.466
Management Range Narrower ($n = 192$)	0.061
t-statistic on difference	3.90#
probability	0.001

Panel C. Instances of Increases vs. Decreases in Analyst Standard Deviation:

	Analyst Standard Deviation		
Relation of Management to Analyst Range	*Increases*	*Decreases*	*Total*
Management Range Wider	20	16	36
Management Range Narrower	55	143	198
Total	75	159	234
Chi-square test of association	10.79		
probability	0.001		

SDDROP = percentage drop in IBES analyst forecast standard deviations from the month prior to the management forecast release to the month following release (see eq. [4] in text); truncated at 100 percent;
SQCPE = standardized, squared three-day market model prediction error (see eq. [8] in text); and
RELWIDE = width of management forecast range relative to analyst forecast range (see eqs. [5] through [7] in text).

* Significant at 0.05 level in one-tailed test.
** Significant at 0.01 level in one-tailed test.
# The hypothesis of equal variance is rejected. Statistic is based on the assumption of unequal variance.

using alternative measures of both variables. In panel B, we simply note whether the management range (*MFWIDE*) is wider or narrower than the analyst forecast range (*AFWIDE*), and compute *SDDROP* for each group.[20] For the 24 instances when managers conveyed a wider range than analysts, mean *SDDROP* is equal to −0.466 which is interpreted as a mean *increase* in standard deviation through time. For the 192 instances when the manager released a narrower range, mean *SDDROP* equals 0.061. The t-statistic on the difference is 3.90 ($p < 0.001$), providing additional evidence in support of alternative hypothesis H2.[21]

---

[20] Two observations can be added to the analysis to yield 216. These observations had *AFWIDE* equal to zero but not equal to *MFWIDE*. Thus, *RELWIDE* could not be computed for the panel A tests.

Rather than calculating the continuous measure, *SDDROP*, panel C repeats the analysis by counting the number of instances where the analyst forecast standard deviation either increased or decreased. Although this technique ignores magnitude information, an additional 18 observations are added. In these cases the analyst forecast standard deviation equalled zero at the management forecast date and *SDDROP* was undefined or *RELWIDE* was undefined. An advantage of this analysis (given that we reject the null) is its bias in favor of the null hypothesis since it reduces the influence of larger negative values of *SDDROP*. Large negative values are caused by small denominators (analyst consensus) and negative numerators (decreases in analyst consensus), and thus represent the instances where an imprecise forecast has the greatest opportunity to change precise priors.

Of the 198 instances when a manager issued a forecast with a narrower width than the analyst forecast range, there were 143 instances of a subsequent decrease in analyst forecast standard deviation (72.2 percent). In the 36 cases when a manager widened the range, only 16 decreases in analyst standard deviation (44.4 percent) occurred. The chi-square statistic of 10.79 ($p < 0.001$) rejects the null hypothesis of no difference in the proportion of standard deviation decreases through time. This result is particularly strong given the tendency for standard deviations to decrease through time.

IV. Summary and Conclusions

This study examines the effects of information precision on equity pricing and on the assessment of earnings uncertainty. Results on the equity pricing issue jointly support (1) the expectations adjustment hypothesis that managers will credibly label their forecasts as to precision, and (2) Kim and Verrecchia's (1991) prediction that the price reaction to the public release of information is a positive function of the information's precision. We are able to measure precision directly because managers explicitly disclose their level of uncertainty in voluntary forecasts.

Results on earnings uncertainty support the predictions of Holthausen and Verrecchia (1990) and Morse et al. (1991). These studies model a decrease in investors' consensus as a positive function of the magnitude of signal surprise and the dispersion of the perceived precision of the signal. Our results corroborate the Morse et al. (1991) finding of a positive relation between signal surprise and increase in consensus analyst forecast variance. Also, managers' explicit labeling of forecasts as uncertain through range disclosure permits the direct calculation of management forecast precision relative to analyst forecast precision, which allows us to document a separate effect of relative management forecast precision on changes in the consensus of beliefs.

---

[21] The hypothesis of equal variance across groups is rejected at the 0.001 significance level. The reported t-statistic is based on the assumption of unequal variance. *SDDROP* is bounded by 1.0 on the positive side by construction (see eq. [4]). We truncated *SDDROP* at -1.0 on the negative side for all tests reported in table 3. We repeated all table 3 tests (1) employing no truncation rule and (2) after truncating at the next most extreme value of *SDDROP* until we were truncating 5 percent of the observations. The conclusions obtained from table 3 do not change based upon any truncation scheme.

Appendix

Alternative Measures of Unexpected Earnings for Imprecise Forecasts

The purpose of this appendix is to examine the robustness of our results to the method of measuring the unexpected earnings conveyed by a management forecast. Two sets of tests are performed. First, equation (1) is reestimated by replacing *UE* with a one (zero) dummy variable for $UE \geq 0$ ($UE < 0$). This variable, denoted *SIGN*, is unambiguously determinable for a subsample of 502 annual forecasts and 111 interim forecasts. Unambiguous determination of *SIGN* occurs in all instances for point forecasts. *SIGN* is also clearly determinable for closed-interval forecasts when *both* endpoints of the interval are either above or below the median analyst forecast expectation, for minimum type forecasts when the minimum is above the expectation, and for maximum type forecasts when the maximum is below the expectation. Table A1 provides the results of estimating equation (1) for the subsample.

The results in table A1 are consistent with the table 2, panel B results with respect to hypothesis H1, and slightly stronger on the coefficient of interest, b_3. Also, the significant intercept shift b_2 is of the opposite sign of the negative and significant b_0 intercept, consistent with the notion of a dampening effect for imprecise forecasts.

In the second set of tests, we biased the design in favor of the null hypothesis by searching for the definition of *UE* for imprecise forecasts that would maximize its association with unexpected returns. For closed-interval forecasts, there are an infinite number of potential forecasts within the interval that could be used to define *UE*, so testing each is impractical. Accordingly, we chose the bottom, middle, and top of the closed-end interval for three potential management forecast measures. We then chose the low, median, and high IBES analyst forecasts as three measures of expectations. We computed *UE* using the nine possible combinations of forecasts and expectations and conducted nine regressions with *CPE* as the dependent variable. Table A2, panel A presents unadjusted R^2 for each regression.

The combination that yields the highest R^2 (0.1344) is the middle of the management forecast range compared to the median analyst forecast as the expectation. This is the method we use in our table 2 tests to measure *UE* for range forecasts.

We followed a similar procedure for open-interval forecasts. Unfortunately, there is no middle or top for minimum type forecasts and no middle or bottom for maximum type forecasts. We created "pseudo-ranges" for these forecasts using the IBES standard deviation. Minimums (maximums) were defined as the bottom (top) of the range. Adding (subtracting) one standard deviation created the middle of the range, and adding (subtracting) two standard deviations created the top (bottom) of the range. Panel B of table A2 provides the R^2 values for

Table A1

Replication of Table 2, Panel B Using Subsample of Forecasts With Unambiguous Measure of Sign of Unexpected Earnings

Model: $CPE_i = b_0 + b_1(SIGN_i) + b_2(IMPRECISE_i) + b_3(IMPRECISE_i{*}SIGN_i) + e_i$.

	Annual (n = 502)		*Interim* (n = 111)	
Coefficient	*Estimate*	*t-statistic*	*Estimate*	*t-statistic*
b_0	−0.034	−5.71**	−0.042	−2.94**
b_1	0.037	4.52**	0.055	2.37**
b_2	0.013	3.46**	0.026	2.56**
b_3	−0.024	−3.82**	−0.049	−2.99**
R^2	0.0410		0.0840	
Adjusted R^2	0.0352		0.0581	

CPE = three-day cumulative market model prediction error (see eq. [2] in text);

SIGN = sign of unexpected earnings conveyed by the management forecast (see eq. [3] in text for computation of *UE*); 1 (0) for $UE \geq (<) 0$; and

IMPRECISE = level of management forecast imprecision equal to 2 (1,0) for open-interval (range, point) forecasts.

* Significant at the 0.05 level in a one-tailed test.

** Significant at the 0.01 level in a one-tailed test.

Table A2
Coefficients of Determination Using Alternative Measures of Unexpected Earnings

Each table entry is the coefficient of determination from:

$$CPE_i = c_0 + c_1(UE_i) + e_i.$$

Panel A. Closed-Interval Forecasts (n = 190):

Measure of Analysts' Expectations	*Bottom*	*Middle*	*Top*
	Measure of Forecasted Earnings		
Lowest Analyst	0.0438	0.0671	0.0441
Median Analyst	0.0467	0.1344	0.0474
Highest Analyst	0.0264	0.0464	0.0287

Panel B. Open-Interval Forecasts (n = 482):

Measure of Analysts' Expectations	*Bottom*	*Middle*	*Top*
	*Measure of Forecasted Earnings**		
Lowest Analyst	0.0038	0.0061	0.0075
Median Analyst	0.0024	0.0067	0.0103
Highest Analyst	0.0003	0.0017	0.0042

CPE = three-day cumulative market model prediction error (see eq. [2] in text); and
UE = unexpected earnings conveyed by the management forecast (see eq. [3] in text).

* Either the "bottom" or "top" (and in all cases, the "middle") of an open-interval range must be constructed. We used IBES standard deviations for this construction. For minimums ("greater than" type forecasts), we added one (two) standard deviations to obtain the middle (top) of the interval. For maximums ("less than" type forecasts), we subtracted the standard deviations to obtain the middle and bottom of the interval.

the nine possible regressions combining the three points in these "pseudo-ranges" with the three potential analyst forecast expectations.

The highest R^2 (0.0103) is obtained by using the top of the pseudo-range and the median analyst forecast to calculate *UE*. For maximum-type forecasts, this does not deviate from the measurement technique in the table 2 regressions. For minimim-type forecasts, however, the top of the pseudo-range is the minimum plus two standard deviations. Since we did not use this technique in our main tests, we replicated table 2 using the new measurement technique for minimums. The results (not reported in a table) are that the regression R^2 values increase slightly, and that coefficients on the imprecision by unexpected earnings interaction (b_3) decrease slightly but remain significantly negative at 0.086 and 0.019 probability levels for annual and quarterly forecasts, respectively.

In summary, it appears that the results on the precision hypothesis H1 are robust to alternative unexpected earnings measurement schemes. Of the two sets of tests provided in this appendix, the first set probably provides the best indication of robustness. Although the calculation of SIGN discards magnitude information, it does so equally across forecast imprecision levels. Further, limiting the sample only to those forecasts for which SIGN is clearly determinable avoids specification of the measurement technique as a function of the forecast imprecision level. While the second set of tests chooses the measurement technique for *UE* that provides the greatest bias in favor of the null, the null is still rejected. It is impractical, however, to consider all measurement alternatives.

References

Ajinkya, B., and M. Gift. 1984. Corporate managers' earnings forecasts and symmetrical adjustment of market expectations. *Journal of Accounting Research* 22 (Autumn): 425–44.

——, R. Atiase, and M. Gift. 1991. Volume of trading and the dispersion in financial analysts' earnings forecasts. *The Accounting Review* 66 (April): 389–401.

Atiase, R. 1985. Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research* 23 (Spring): 21–36.

Baginski, S. P., and J. M. Hassell. 1990. The market interpretation of management earnings forecasts as a predictor of subsequent financial analyst forecast revision. *The Accounting Review* 65 (January): 175–90.

Beaver, W. 1968. The information content of annual earnings announcements. *Journal of Accounting Research* 6 (Supplement): 67–92.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics*. New York: John Wiley.

Brown, P., G. Foster, and E. Noreen. 1985. *Security Analyst Multi-Year Earnings Forecasts and the Capital Market*. AAA Studies in Accounting Research #21.

Cox, C. 1985. Further evidence on the representativeness of management earnings forecasts. *The Accounting Review* 60 (October): 692–701.

Diamond, D. 1985. Optimal release of information by firms. *Journal of Finance* 40 (September): 1071–94.

Hassell, J. M., and R. Jennings. 1986. Relative forecast accuracy and timing of earnings forecast announcements. *The Accounting Review* 61 (January): 58–75.

Holthausen, R., and R. Verrecchia. 1990. The effect of informedness and consensus on price and volume behavior. *The Accounting Review* 65 (January): 191–208.

Jennings, R. 1987. Unsystematic security price movements, management earnings forecasts, and revisions in consensus analyst earnings forecasts. *Journal of Accounting Research* 25 (Spring): 90–110.

Kim, O., and R. Verrecchia. 1991. Trading volume and price reactions to public announcements. Journal of Accounting Research 29 (Autumn): 302–21.

King, R., G. Pownall, and G. Waymire. 1990. Expectations adjustment via timely management forecasts: Review, synthesis, and suggestions for future research. *Journal of Accounting Literature* 9: 113–44.

McNichols, M. 1989. Evidence of informational asymmetries from management earnings forecasts and stock returns. *The Accounting Review* 64 (January): 1–27.

Morse, D., J. Stephan, and E. Stice. 1991. Earnings announcements and the convergence (or divergence) of beliefs. *The Accounting Review* 66 (April): 376–88.

Patell, J. 1976. Corporate forecasts of earnings per share and stock price behavior: Empirical tests. *Journal of Accounting Research* 14 (Autumn): 246–76.

Penman, S. 1980. An empirical investigation of the voluntary disclosure of corporate earnings forecasts. *Journal of Accounting Research* 18 (Spring): 132–60.

Pownall, G., C. Wasley, and G. Waymire. 1993. The stock price effects of alternative types of management earnings forecasts. *The Accounting Review* 68 (October): 896–912.

Waymire, G. 1984. Additional evidence on the information content of management earnings forecasts. *Journal of Accounting Research* 22 (Autumn): 703–18.

Ziebart, D. 1990. The association between consensus of beliefs and trading activity surrounding earnings announcements. *The Accounting Review* 65 (April): 477–88.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 928–941

Earnings-Based Compensation Plans, Performance, and Capital Expenditure Policy in the Motor Carrier Industry

Charles R. Enis
The Pennsylvania State University

SYNOPSIS AND INTRODUCTION: This research examines the effect that adoption of earnings-based incentive plans has on performance and capital expenditure policy in the motor carrier industry. The motor carrier industry provides the opportunity to study the effect of incentive plans on small, closely-held organizations. Restricting the sample to a single industry controls for variations in accounting procedures, heterogeneous production functions, and other environmental factors.

Firms that adopted bonus and performance plans are matched with similar firms without such plans. The matching criteria are size, carrier type, and freight classification. The dependent variables are operating ratio, net income, capital expenditures, and maintenance expenditures.

The analysis investigates the extent to which the groups differ in performance and investment policy before and after plan adoption. The results indicate that motor carriers that adopted bonus plans had out-performed the respective control firms during the post-adoptive period. This relatively improved performance is not linked to reductions in capital or maintenance expenditures. Although the results for performance plan adopters are in the expected directions, they are not statistically significant.

Key Words: *Incentive compensation plans, Agency problems, Capital expenditure policy, Earnings-based compensation.*

Data Availability: *The data are available from the author upon request.*

The helpful comments of the following are sincerely appreciated: Chee Chow, Herb Hunt, Ed Morash, Jim Waegelein, Jim McKeown, John Fellingham, three anonymous reviewers, and the participants in accounting research workshops at University of Nebraska and Virginia Polytechnic and State University. Also greatly appreciated is the research, data collection, transcription, and proofing assistance of Cheryl Burgess, Dave Slockbower, Michelle Scarano, John Russell, Patrick Cronin, Rebecca DaSilva, John Savage, Louise Talbert, and Doris Lewis. Remaining errors and omissions are the author's responsibility.

Submitted November 1989 and April 1992.
Accepted March 1993.

THIS article consists of four sections. In the background section, I review other studies concerning the effects of incentive plans and also discuss the relevance of the motor carrier industry to the research objectives of this study. The method section presents details concerning the data set, sample selection, matching procedure, and dependent variables. The results section contains the empirical findings, and the conclusion contains a summary and some final comments.

I. Background

Bonus Plans

Bonus plans that pay executives a portion of the firm's annual earnings are intended to increase managerial effort by allowing managers to share in the profits resulting from their efforts, and thus better align the interests of shareholders and managers (Shavell 1979). Prior studies have linked bonus plan adoption with changes in managerial behavior. Healy (1985) provides some evidence that bonus plans induce executives to manage earnings to favor their incentive awards. Tehranian and Waegelein (1985) observe a positive market reaction and positive unexpected earnings associated with the announcement of the adoption of a bonus plan ten months earlier. Larcker (1987) finds that the adoption of bonus plans motivates managers to supply effort more in line with shareholder objectives.

This study examines the extent to which the findings of prior research on large publicly traded firms can be extended to smaller closely-held firms, where agency issues presumably differ. The motor carrier industry is examined because it is comprised of predominately small, nonpublicly traded firms that are required to report substantial financial and nonfinancial information to the Interstate Commerce Commission (ICC). In this setting, principals are small ownership groups that often participate in decision making, and agents are managers outside of these groups.

In prior studies of incentive plans, accounting and share price data have provided consistent measures of firm performance (Ely 1991; Kumar and Sopariwala 1992; Lambert and Larcker 1987). Because firms in this study are not publicly traded, two accounting measures of performance are used: net income and operating ratio. The latter measure is the ratio of operating expenses to operating revenues and is an inverse indicator of managerial efficiency.

This study offers two possible expected relations between bonus plan adoption and changes in firm performance. The first is based on the presumption that an owner of a closely-held business has a larger stake in the firm than an owner of a publicly traded portfolio has in any given stock. This presumption suggests that a shareholder of a closely-held firm is likely to be less risk-neutral than a shareholder of a publicly traded firm. Arya (1991) shows that the less risk-neutral the principal is, the more likely risk will be shared with the agent through outcome-contingent compensation such as a bonus plan. The motivation to adopt a bonus plan is to align the interest of the owner and manager, and thus induce greater effort from managers to decrease the firm's operating ratio and increase net income.

The second possibility is based on the presumption that managers of closely-held firms encounter greater monitoring than managers of publicly traded firms (Kaplan 1989; Smith 1990). If such monitoring is sufficient to assure that managers supply desired levels of effort, then small firms that adopt bonus plans do so for reasons other than incentive effects, for example, to minimize the joint tax burdens of firms and managers (Miller and Scholes 1982). Thus, no association is expected between bonus plan adoption and the performance measures.

This empirical work is aimed at determining which of these two opposing possibilities is dominant. There is a paucity of empirical evidence concerning the association between bonus plan adoption and subsequent firm performance, particularly for closely-held firms. Thus, the first hypothesis is:

H1: The operating ratios (net incomes) for bonus plan adopters are smaller (greater) than those for the control firms.

Performance Plans

Prior studies of the effect of performance plans and other long-range incentive contracts on managerial behavior have been mixed.[1] Some studies have documented an association between long-range plans and superior share price performance (Brickley et al. 1985; Larcker 1983; Tehranian et al. 1987a, 1987b). In contrast, Gaver et al. (1992) did not find that performance plan adoptions significantly increase shareholder wealth. These researchers cite timing problems in analyzing stock market reactions to information disclosed in proxy statements to explain the conflict between their findings and those of earlier studies.

Although most studies of performance plans involve market measures of performance, Kumar and Sopariwala (1992) did find significantly greater growth in earnings per share for plan adopters relative to control firms in the second year following performance plan adoption. Because of linkages in the empirical literature between long-range plan adoptions and market and accounting measures of performance, expectations regarding performance plan adoptions and profitability measures mirror those discussed for bonus plans. Thus, the present study tests the following hypothesis:

H2: The operating ratios (net income) for performance plan adopters are smaller (greater) than those for the control firms.

Performance plans are also examined to investigate their role in mitigating the horizon problem. In the absence of equity-based compensation, managers' decision-making horizons are limited by their tenure with the firm. This horizon is even shorter if bonus plans induce managers to focus on annual profitability. Performance plans, whereby managers are compensated as a function of earnings over periods longer than a year (usually three to six years) are intended to mitigate the horizon problem by motivating managers to focus more on long-run profitability (Larcker 1983; Shavell 1979; Smith and Watts 1982). Behavior detrimental to the long-run profitability of the firm, such as reducing capital expenditures and maintenance expenditures, is consistent with the horizon problem.

Performance plans are suggested to counter the alleged tendency of bonus plans to discourage capital expenditures. Smith and Watts (1982) assert that bonus plans motivate the rejection of positive net present value projects with long investment horizons, especially those with substantial cash flows late in the life of the project. Larcker (1983) argues that performance plans may result in increased capital expenditures by attracting managers to such projects. He also finds a positive association between the adoption of performance plans and growth in capital expenditures in his study of 25 matched pairs of COMPUSTAT firms from ten industries. Lewellen et al. (1987) conclude that long-range compensation plans are effective in increasing capital expenditures by motivating managers to undertake risky investments that will enhance shareholder wealth. Thus, a negative (positive) relation is expected between bonus (performance) plan adoption and capital expeditures; however, Waegelein (1988) finds the opposite regarding bonus plan adoption. The present study tests the following hypotheses:

---

[1] Prior studies have included stock-based plans, such as stock options, stock appreciation rights, phantom stock, dividend units, etc., along with performance plans in examining the effects of long-range compensation plans (Brickley et al. 1985; Tehranian et al. 1987a, 1987b). Performance plans are often used as long-range incentive plans when stock-based plans are not practicable (Gaver 1992). Because the firms in the present study are predominately closely held carriers, stock-based features are uncommon. Of 307 respondents, only 12 report stock-based plans (see table 1). Thus, the present study lumps stock-based plans with performance plans. Brickley et al. (1985) did not detect differential effects across types of long-range plans.

H3: Capital expenditures for bonus plan adopters are smaller than those for the control firms.
H4: Capital expenditures for performance plan adopters are greater than those for the control firms.

Unlike capital expenditures that are expensed over several periods, maintenance expenditures are expensed as incurred. Thus, a reduction in the latter could improve short-run earnings, and hence remuneration from a bonus plan. Although no study has specifically examined maintenance expenditures, Larcker (1987) observes a negative relation between bonus plan adoption and discretionary expenditures, but Waegelein (1988) finds no significant relation between plan adoption and either advertising or research and development expenditures. The present study tests the following hypotheses:

H5: Maintenance expenditures for bonus plan adopters are smaller than those for the control firms.
H6: Maintenance expenditures for performance plan adopters are greater than those for the control firms.

Role of the Motor Carrier Industry

In addition to the advantages of using smaller, closely-held firms for analysis of bonus and performance plans, the motor carrier industry has other interesting features. Managers of trucking firms do not have as much freedom to manipulate reported earnings as managers in other industries. To begin with, the ICC requires uniform accounting methods for the consistent asset mix found in the industry.[2] Inventories and manufacturing processes are not factors in trucking operations. Thus, managers cannot manage earnings by adjusting inventory levels or manufacturing overhead allocations. Such uniformity and consistency control for confounding variables that have been suggested as influencing the relation between executive compensation and performance measures (Healy et al. 1987; Lewellen et al. 1987). Although uniform accounting rules are no guarantee that earnings management does not occur in the motor carrier industry, such standardization offers a greater level of confidence that the findings in this study are more attributable to decisions that alter the real value of the firm rather than to earnings manipulations.

---

[2] All class I and II motor carriers are required by law to base their financial reporting on "The Uniform System of Accounts" (Title 49 CFR §1249.2). For discussions and details related to the accounting procedures that are required of trucking firms, see American Trucking Associations (1991) and KPMG Peat Marwick (1991).

Lanen and Larcker (1992) find a modest association between production efficiency, changes in the regulatory environment and the adoption of compensation contracts in the electric utility industry. Similarly, a major regulatory change occurred when the Motor Carrier Act of 1980 reduced market restrictions on trucking firms. This law gave motor carriers greater freedom in setting rates, selecting routes, and entering new markets (Harper 1980; ICC 1981). The matched-sample design in this study controls for changes in the regulatory environment by comparing adopting and nonadopting carriers across event years with the same regulatory regimes.[3] Regulation is a source of monitoring absent in nonregulated industries. Acceptance of the hypotheses are less likely if regulatory restrictions on managerial decision making attenuate the need for incentive compensation (Ely 1991).

II. Method

General Description

Differences in performance and investment policy measures between carriers in the experimental groups and their respective control firms are tested for statistical significance by using the Wilcoxon matched-pairs signed ranks test (hereafter Wilcoxon). This test is applied to the plan adoption year (year 0), each of the four preceding years (years $-4,\ldots,-1$), and each of the four subsequent years (years $+1,\ldots,+4$). Because adoption years vary across experimental firms, year-specific events are diversified away. To test whether values of the dependent variables change from the pre-adoption period (years $-4,\ldots,-1$) to the post-adoption period (years $+1,\ldots,+4$), the sum of each dependent variable (Y_i) over the four-year post-adoption period for the experimental (Y_{iEA}) and control (Y_{iCA}) carriers is first computed. Similar figures are then determined for the experimental and control firms over the pre-adoption period; (Y_{iEB}) and (Y_{iCB}), respectively. The Wilcoxon test is applied to ascertain whether ($Y_{iEA}-Y_{iCA}$) is statistically different from ($Y_{iEB}-Y_{iCB}$) for each dependent variable i. The median difference of the contrast $[(Y_{iEA}-Y_{iCA})-(Y_{iEB}-Y_{iCB})]$ measures the magnitude and direction of changes in the real values of the dependent variables between the pre- and post-adoption periods (see Larcker 1987; Waegelein 1988).

As additional analysis of the relation between performance and investment policy before and after plan adoption, percentage changes in the dependent variables (PCY_i) from the pre- to the post-adoption period are computed for the experimental firms; $PCY_{iE}=(Y_{iEA}-Y_{iEB})/Y_{iEB}$. Next, Spearman correlations between the PCY_{iE} figures for the performance and investment policy variables are tested for statistical significance (Kumar and Sopariwala 1992).

---

[3] Regulation can also vary across jurisdictions and carrier types; for example, speed, size and weight limits are more restrictive in the more congested eastern states than in the western states. Hazardous material carriers are more regulated than others. This study attempts to control for such differences by matching on carrier type and freight.

Table 1
Survey Respondents According to Size and Compensation Plan

Size[a]	*Compensation Plan Category*			
	Performance Plan	*Bonus Plan*	*No Plan*	*Row Totals*
OR > \$75 million	13 (28%)	37 (30%)	16 (11%)	66
\$15 million ≤ *OR* ≤ \$75 million	17 (37%)	48 (40%)	46 (33%)	111
OR < \$15 million	16 (35%)	36 (30%)	78 (56%)	130
Column Totals	46 (100%)	121 (100%)	140 (100%)	307[b]

Note: The null hypothesis that size and compensation plan categories are unrelated is rejected, $X^2=24.1$, d.f.$=4$, $p<0.001$.

[a] Operating revenue (OR) in 1967 dollars.

[b] Because the firms that received questionnaires were randomly selected, some publicly traded firms appear in this total.

Data and Sample Selection

Compensation Plans. Data for the matching and dependent variables are obtained from the *TRINCs Blue Book of the Trucking Industry* (hereafter TRINCs). The Consumer Price Index is used to restate all values to 1967 dollars. A direct mail survey of 1,500 motor carriers randomly selected from TRINCs is used to determine whether a firm adopted a plan, and if so, the type of plan and the year of adoption.[4]

Respondent firms are classified as having bonus plans if awards are based solely upon some annual measure of earnings. A performance plan classification is made if the plan includes a provision whereby an award is based on earnings variables measured over several years.[5] Of the 307 respondents (table 1), 46 had adopted performance plans, and 121 had bonus plans.

Matching. With the application of four restrictions, 50 of the 121 carriers that adopted bonus plans and 30 of the 46 carriers that adopted performance plans are paired with carriers having no incentive plans. The first restriction is that firms must have conducted operations for five years before adopting an incentive plan during 1963–1983.[6] This requirement is necessary because of limitations associated with using

---

[4] The classification of incentive plans for publicly traded carriers is verified by inspecting proxy statements filed with the Securities and Exchange Commission (SEC). For nonpublicly traded carriers data were gathered from the survey.

[5] For example, a major household goods carrier offers a performance plan whereby the annual contingent awards are modified by final awards based on a formula for weighing the annually compounded growth rate during the subsequent three-year performance period. Earnings per share, return on average shareholder equity, and return on average net capital are the elements in this formulae, with each measurement factor considered independently. The company pays its awards in the fourth year after the contingent award. It requires the participant to remain employed during the performance period. A performance plan classification does not rule out rewards for yearly performance. All performance plan adopters used in the sample offer such a provision.

[6] Event year −5 is required to compute capital expenditures for event year −4; 26 (6) bonus (performance) plan adopters were eliminated from the final sample because they lacked this pretest period; see item a, table 3.

Table 2

Summary Statistics of the Experimental and Control Carriers Regarding Size

	Mean		*Median*		
Size Variable[a]	*Experimental*	*Control*	*Experimental*	*Control*	*Wilcoxon* |*Z*|[b]
Bonus Plan					
Operating Revenues ($n = 50$)	8,645	6,283	2,898	3,506	1.09
Revenue Equipment[c] at Cost ($n = 50$)	3,362	2,437	890	1,389	0.20
Performance Plan					
Operating Revenues ($n = 30$)	25,702	22,884	4,881	3,041	1.61
Revenue Equipment[c] at Cost ($n = 28$)	6,575	5,050	1,285	1,300	0.34

[a] One year prior to the plan adoption year, in constant thousands of 1967 dollars.

[b] Absolute value of the Wilcoxon Z; none is significant at the 0.10 level, two-tail.

[c] The cost of trucks, tractors, and trailers owned by the carrier.

the TRINCs data base.[7] The remaining restrictions arise from the three matching criteria.

The first matching criterion is size, as measured by operating revenues for the year prior to the adoption year. A minimum condition for a match is that the ratio of the smaller firm's operating revenues to the larger firm's be at least 0.2. This ratio is greater than 0.5 for 61.2 percent of the final pairings. Table 2 presents summary statistics showing that the experimental and control firms do not differ significantly in size. However, firms that adopted performance plans tend to be larger than bonus plan adopters ($t=1.83$, $p=0.07$, two-tail). Table 3 shows the years in which plans were adopted and the attrition from the survey respondents to the final matched pairs.

The second requirement is that both carriers be of the same type. Contract, local freight, and consolidated carriers are paired with carriers having similar legal restrictions. Third, the pairs must have the same freight classification: general freight, household goods, refrigerated and agricultural products, petroleum products, and other specialized freight. This criterion is used so that paired firms maintain and invest in similar revenue equipment (trucks, tractors, and trailers) and serve shippers in similar sectors of the economy.

Dependent Variables

Performance Measures. The two performance measures examined in this study are the operating ratio and net income. The operating ratio is closely followed by trucking

---

[7] The TRINCs data base consists of a series of yearbooks stored in subscribers' libraries in bound hardcopy or microfilm form. The data are not publicly available on computer tape. Dun and Bradstreet disbanded TRINCs in 1985. As a result, 1983 is the last reporting year for which TRINCs published a yearbook. In the present study, data for years after 1983 were obtained from the *Motor Carrier Annual Report* published by the American Trucking Associations.

The earliest TRINCs yearbook available from libraries reports data for 1958. Thus, if 1958 is event year −5, 1963 is the adoption year 0. Because of this constraint, carriers that adopted plans prior to 1963 were excluded from the study.

Table 3
Number of Plan Adoptions by Year

Year	Number of Plan Adoptions: Bonus Plan	Number of Plan Adoptions: Performance Plan
1963	1	1
1967	1	0
1968	3	0
1970	4	1
1972	3	1
1973	4	2
1974	4	3
1975	4	2
1976	6	1
1977	3	1
1978	7	5
1979	4	1
1980	4	3
1981	2	5
1982	0	2
1983	0	2
Total Included	50	30
Reasons for Exclusion		
No pretest period[a]	26	6
Insufficient data[b]	16	5
No match[c]	11	0
Adoption not timely[d]	18	5
Total Excluded	71	16
Total Respondents	121	46

[a] The firm did not conduct operations in years prior to plan adoption.

[b] Insufficient data were reported in TRINCs for any meaningful analysis within the purview of the present empirical work.

[c] A control firm could not be found that satisfied the matching criteria for the pairing.

[d] The adoption date fell outside the 1963–1983 period.

industry analysts and is frequently used in bonus plans in the motor carrier industry (GAO 1991; Lisciandro 1983). The operating ratio is a surrogate for unobservable managerial effort. Net income is examined because all bonus and performance plans in this study are based on one or more variants of earnings measured over one or several years.

Capital Investment Policy. The relation between incentive compensation and investment policy is addressed by analyzing capital and maintenance expenditures on revenue equipment. This asset group is of interest because the expansion, replacement, and maintenance of truck fleets is often governed by some systematic managerial policy that could be influenced by changes in executive compensation packages (Frye 1980).

Capital expenditures on revenue equipment for year t are obtained by computing the difference between the amount reported in the "Revenue Equipment Owned at Cost" account for years t and $t-1$.[8] This measure is an indication of the extent to which a motor carrier is expanding or reducing its fleet capability and increasing or decreasing the replacement cycle of its revenue equipment.

III. Results

The median and median differences in operating ratio (a proxy for effort) and net income and the associated Wilcoxon $|Z|$ statistics are presented in table 4. This table also reports the contrast between the pre- and post-adoption periods. Table 5 is organized in the same manner except that the dependent variables are capital expenditures and maintenance expenditures.[9]

If adoptions of incentive plans influence managers to supply greater effort, which leads to improved profitability, one would expect no significant differences in performance measures between experimental and control carriers for years $-4,\ldots,-1$. But one would expect the operating ratio (net income) numbers for the experimental carriers to be lower (greater) than those of the respective control carriers for years $+1,\ldots,+4$.

As shown in table 4, the median operating ratio for the bonus plan adopters is significantly lower than that of firms with no plan for year +4. The increase in median operating ratios is slower for the experimental firms. Thus, the greater managerial efficiency associated with bonus plan adopters in year +4 is relative to their respective control firms rather than to their past performances. Further evidence of the increasing disparity between effort supplied by managers of experimental and control firms is the highly significant Wilcoxon $|Z|$ statistic of 3.09 associated with contrasting differences in operating ratios for the pre- and post-adoption periods. In contrast to bonus plans, the adoption of performance plans has little influence on operating ratios.

According to table 4, the adoption of bonus plans seems to be associated with increased net income. No significant differences are observed between the net income numbers of the bonus plan firms and the respective control firms during any of the pre-adoption years. Also, the contrast between the pre- and post-adoption periods links bonus plan adoption with superior earnings performance. In contrast, table 5 shows no significant difference between net income and performance plan adoption.[10]

With respect to capital expenditures and maintenance expenditures, the bonus (performance) plan adopters show no significant difference from control firms. This finding is contrary to the assertion that incentive plan adoptions are associated with horizon problems in the motor carrier industry.

---

[8] The "Revenue Equipment Owned at Cost" account as specified by the "Uniform System of Accounts" is original historical cost of revenue equipment in service and is not reduced by accumulated depreciation.

[9] The variation in sample size reported in the tables resulted from some carriers not reporting a full set of data for certain years. The two samples do not overlap.

[10] The analysis reported in table 4 was also performed by using profit margin and return on equity as profitability measures. The results are similar in that experimental firms are more profitable than their respective control firms after adoption, and this relation is stronger for the bonus plan sample.

Table 4
Summary Statistics of Differences in Effort and Earnings Performance by Carriers

	Operating Ratio[a]					Net Income				
		Medians					Medians			
	No. of Pairs	*Exper.*	*Control*	*Differ.*[b]	*Wilcoxon* |Z|	No. of Pairs	*Exper.*	*Control*	*Differ.*[b]	*Wilcoxon* |Z|
Bonus Plan Adoption										
Year −4	50	95.6	94.9	1.5	1.81	50	62	92	−11	0.89
Year −3	50	95.5	95.2	0.0	1.19	50	64	85	−3	0.33
Year −2	50	95.8	94.3	1.1	1.34	50	72	115	2	0.55
Year −1	50	95.9	94.6	0.7	0.82	50	78	112	7	0.31
Year 0	50	95.7	95.0	−0.2	0.09	50	70	97	25	1.50
Year +1	50	96.2	95.6	−0.5	0.71	50	54	107	40	2.23*
Year +2	50	96.8	96.7	−0.9	0.88	50	70	45	9	1.60
Year +3	50	96.3	97.0	−0.2	0.61	50	77	55	39	2.35**
Year +4	50	96.2	97.8	−2.4	2.36**	50	79	63	54	1.95*
Contrast[c]	50			−11.1	3.09**	50			98	1.80*
Performance Plan Adoption										
Year −4	28	96.7	94.7	1.0	1.21	28	83	99	−14	0.72
Year −3	28	94.8	93.9	2.0	0.86	28	137	147	−6	0.16
Year −2	28	96.2	94.6	2.8	1.30	28	98	93	4	0.17
Year −1	30	96.9	94.1	2.0	1.53	30	120	115	−23	1.02
Year 0	30	97.4	94.1	1.0	1.63	30	78	135	14	0.17
Year +1	30	98.0	95.2	0.4	1.34	30	109	111	−11	0.28
Year +2	30	97.8	96.7	−0.4	0.35	30	90	82	18	0.29
Year +3	29	97.5	96.6	1.7	0.45	29	128	105	9	0.38
Year +4	28	98.5	97.5	1.0	0.48	28	135	98	4	0.20
Contrast[c]	30			−7.3	0.68	30			32	0.22

[a] Operating Ratio=operating expenses as a percentage of revenues. Excluded are extraordinary items, income taxes, and interest. This profitability ratio is used in the trucking industry as an inverse measure of managerial efficiency, and is a surrogate for managerial effort in this study.

[b] Differences are calculated by subtracting the data for the control carriers from the data for the experimental carriers. The operating ratio is expressed in percentage points and net income in thousands of 1967 dollars.

[c] The contrast between the relative pre- and post-adoption levels of managerial efficiency and net income.

* Statistically significant at $p \leq 0.05$, one-tail.

** Statistically significant at $p \leq 0.01$, one-tail.

Table 5
Summary Statistics Comparing Differences in Capital Expenditures and Maintenance Expenditures

	Capital Expenditures[a]					Maintenance Expenditures				
		Medians					Medians			
	No. of Pairs	*Exper.*	*Control*	*Differ.*[b]	*Wilcoxon* |*Z*|	*No. of Pairs*	*Exper.*	*Control*	*Differ.*[b]	*Wilcoxon* |*Z*|
Bonus Plan Adoption										
Year –4	45	8	46	–6	0.02	50	179	291	–72	0.10
Year –3	45	2	7	20	1.53	50	211	297	–56	0.06
Year –2	50	11	7	26	1.07	50	243	276	–25	0.04
Year –1	50	–1	–11	49	0.55	50	234	293	–38	0.05
Year 0	50	11	8	8	1.40	50	230	291	–43	0.41
Year +1	50	7	2	25	0.04	50	244	296	–22	0.57
Year +2	50	8	–28	88	2.44	50	255	229	–15	0.80
Year +3	50	9	–11	22	1.05	50	272	275	5	1.07
Year +4	50	16	–10	–10	0.26	50	246	259	55	1.58
Contrast[c]	50			–29	0.82	50			72	0.14
Performance Plan Adoption										
Year –4	27	42	95	–5	0.47	28	286	240	33	1.29
Year –3	28	4	21	–18	1.14	28	252	257	20	1.02
Year –2	28	35	8	–14	0.11	28	299	268	25	0.83
Year –1	26	19	18	–58	0.76	29	322	271	1	0.38
Year 0	27	28	19	–10	1.00	29	318	248	27	0.34
Year +1	28	–31	–19	–24	1.25	30	321	266	–9	0.36
Year +2	29	37	5	23	0.14	30	311	290	30	0.11
Year +3	29	24	14	41	0.18	29	371	359	38	0.64
Year +4	28	–6	35	–20	0.18	28	379	358	7	0.14
Contrast[c]	28			–39	0.50	30			–52	0.79

[a] Difference in revenue equipment at cost (i.e., before depreciation) from that of preceding year.

[b] Differences are calculated by subtracting the data for the control carriers from the data for the experimental carriers. Capital and maintenance expenditures are expressed in thousands of 1967 dollars.

[c] Contrast refers to the contrast between the relative pre- and post-adoption levels of capital and maintenance expenditures.

* Statistically significant at $p \leq 0.05$, one-tail.

** Statistically significant at $p \leq 0.01$, one-tail.

Table 6

Spearman Correlations Relating Percentage Changes in Performance and Investment Policy Variables Before and After Plan Adoption for Experimental Firms

	Investment Policy Variables			
Performance Variables	*N*[a]	*Capital Expenditures*	*N*	*Maintenance Expenditures*
Bonus Plan Adoption				
Operating Ratio	50	−0.371[b] (0.004)	50	−0.294 (0.019)
Net Income	47	0.499 (0.001)	47	0.412 (0.002)
Performance Plan Adoption				
Operating Ratio	23	−0.127 (0.283)	25	−0.013 (0.475)
Net Income	21	0.716 (0.001)	23	0.334 (0.060)

[a] Number of firms for which meaningful percentage changes in the dependent variable from the pre- to the post-adoption period could be computed.

[b] Relevant Spearman correlation coefficients; numbers in parentheses represent the relevant *p*-values (two-tailed).

Table 6 shows a significant (negative) positive correlation between percentage change in (operating ratio) net income and each investment policy variable from before to after bonus plan adoption. However, the only such significant ($p<0.05$) relation regarding performance plan adoption is the positive correlation between net income and capital expenditures. These findings do not support the notion that increased managerial effort and profitability following bonus plan adoption are connected with reductions in capital expenditures and maintenance expenditures.

IV. Conclusions

This study suggests that real incentive effects are associated with bonus plan adoptions by motor carriers; these are firms that are predominately small, regulated, and closely held, and therefore principal-agent conflicts are expected to be less evident. Contrary to such reasoning, the results link bonus plan adoption with an apparent greater managerial effort and profitability. This is evident in reduced operating ratios and increased net incomes for experimental firms relative to control firms in the post-adoption period. In short, the greater monitoring encountered in private firms does not totally negate the opportunity to create managerial incentives with bonus plans. However, the findings failed to detect any relation between bonus plan adoptions and horizon problems, as no significant decreases in capital expenditures or maintenance expenditures were observed for bonus plan adopters relative to control firms. These results suggest that monitoring in the trucking industry may be adequate in controlling the horizon problem but not the effort problem.

Unlike bonus plan adoptions, no incentive effects were associated with performance plan adoptions, suggesting that the agency problems might play a little role in adopting these plans. That is, trucking firms may adopt performance plans to save taxes or reduce managerial turnover, as opposed to solving effort and horizon problems.

These results may not be extended beyond the motor carrier industry. More research is needed on managerial incentives and agency problems within the domain of closely-held enterprises.

References

American Trucking Associations. 1991. *Accounting Service for Motor Carriers*. Alexandria, VA: ATA.

Arya, A. 1991. Contracting in a risk-averse principal-agent setting. In *Three Essays on the Principal-agent Model*. Ph.D. dissertation, University of Iowa, Iowa City.

Brickley, J. A., S. Bhagat, and R. C. Lease. 1985. The impact of long-range managerial compensation plans on shareholder wealth. *Journal of Accounting & Economics* 7 (April): 115–29.

Ely, K. M. 1991. Interindustry differences in the relation between compensation and firm performance variables. *Journal of Accounting Research* 29 (Spring): 37–58.

Frye, W. 1980. The optimum replacement point for power-equipment—An integrated approach. In *The NAFC 1980 Award Winning Papers*. Alexandria, VA: ATA.

Gaver, J. J. 1992. Incentive effects and managerial compensation contracts: A study of performance plan adoptions. *Journal of Accounting, Auditing & Finance* 7 (Spring): 137–60.

——, K. M. Gaver, and G. P. Battistel. 1992. The stock market reaction to performance plan adoptions. *The Accounting Review* 67 (January): 172–82.

General Accounting Office. 1991. Report to Congressional Requesters. *Freight Trucking: Promising Approach for Predicting Carriers' Safety Risks*. Washington, D.C.: GAO, Resources, Community and Economic Development Division.

Harper, D. V. 1980. The Federal Motor Carrier Act of 1980: Review and analysis. *Transportation Journal* 20 (Winter): 6–7.

Healy, P. M. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting & Economics* 7 (April): 85–107.

——, S. Kang, and K. G. Palepu. 1987. The effect of accounting procedure changes on CEO's cash salary and bonus compensation. *Journal of Accounting & Economics* 9 (April): 9–34.

Interstate Commerce Commission. 1981. *The Effect of Regulatory Reform on the Trucking Industry: Structure, Conduct, and Performance*. Washington, D.C.: Office of Policy and Analysis.

Kaplan, S. 1989. The effects of management buyouts on operating performance and value. *Journal of Financial Economics* 24 (October): 217–54.

KPMG Peat Marwick. 1991. *Principles and Presentation Series, 1991 Edition: Road Transportation*. Miami, FL: KPMG Peat Marwick.

Kumar, R., and P. R. Sopariwala. 1992. The effect of adoption of long-term performance plans on stock prices and accounting numbers. *Journal of Financial and Quantitative Analysis* 27 (December): 561–73.

Lambert, R. A., and D. F. Larcker. 1987. An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research* 25 (Supplement): 85–125.

Lanen, W. N., and D. F. Larcker. 1992. Executive compensation contract adoption in the electric utility industry. *Journal of Accounting Research* 30 (Spring): 70–93.

Larcker, D. F. 1983. The association between performance plan adoption and corporate capital investment. *Journal of Accounting & Economics* 5 (April): 3–30.

——. 1987. Short-term compensation contracts and executive expenditure decisions: The case of commerical banks. *Journal of Financial and Quantitative Analysis* 22 (March): 33–50.

Lewellen, W., C. Loderer, and K. Martin. 1987. Executive compensation and executive incentive problems: An empirical analysis. *Journal of Accounting & Economics* 9 (December): 287–310.

Lisciandro, P. 1983. *Financial Analysis of the Motor Carrier Industry*. Alexandria, VA: ATA.

Miller, M. H., and M. S. Scholes. 1982. Executive compensation, taxes and incentives. In *Financial Economics: Essays in Honor of Paul Cootner*, edited by W. F. Sharpe and C. M. Cootner. Englewood Cliffs, NJ: Prentice-Hall.

Shavell, S. 1979. Risk sharing and incentives in the principal and agent relationship. *The Bell Journal of Economics* 10 (Spring): 55–73.

Smith, A. J. 1990. Corporate ownership structure and performance: The case of management buyouts. *Journal of Financial Economics* 27 (September): 143–64.

Smith, C., and R. Watts. 1982. Incentive and tax effects of executive compensation plans. *Australian Journal of Management* 7 (December): 139–57.

Tehranian, H., and J. F. Waegelein. 1985. Market reaction to short-term executive compensation plan adoption. *Journal of Accounting & Economics* 7 (April): 131–44.

———, N. G. Travlos, and J. F. Waegelein. 1987a. The effect of long-term performance plans on corporate sell-off-induced abnormal returns. *The Journal of Finance* 42 (September): 933–42.

———, ———, and ———. 1987b. Management compensation contracts and merger-induced abnormal returns. *Journal of Accounting Research* 25 (Supplement): 51–76.

Waegelein, J. F. 1988. The association between the adoption of short-term bonus plans and corporate expenditures. *Journal of Accounting and Public Policy* 7 (Spring): 43–63.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 942–952

COMMENTS

Expectation Formation and Financial Ratio Adjustment Processes: A Comment and An Extension

Samuel Tung
University of Otago (New Zealand)
and National Taiwan University
John Crowe
University of Otago

SYNOPSIS AND INTRODUCTION: Lev (1969) uses a partial adjustment model to examine whether firms adjust their financial ratios to industry averages. Lee and Wu (1988) extend Lev's model by introducing adaptive expectations into the partial adjustment model to examine whether firms adjust their desired target financial ratios. Lee and Wu examine six ratios of 112 U.S. firms from eight industries for the period 1965–1984. Their primary results show that adjustments to target ratios are not instantaneous.

In this comment, we show that Lee and Wu's (1988) use of the Gauss-Newton nonlinear regression method to estimate the structural parameters of the autoregressive form of their partial adjustment-adaptive expectations model is inappropriate. The problem is ambiguity in their estimates, because the two coefficients (the speed of adjustment and the reaction of expectations) in the model occur symmetrically (Maddala 1988, 353). Consequently, Lee and Wu's parameter estimates are sensitive to the selection of the initial parameter values used in their estimation. Furthermore, for the six financial ratios Lee and Wu examine, about 10 to 25 percent of their estimates of the coefficients for the speed of adjustment and approximately

We are thankful to two anonymous referees for their comments. This research benefitted from our discussions with Erkin Bairam and Arnold Zellner. We also acknowledge the computing assistance of Nicola Sewell and financial support from the Otago Research Grant Program.

Submitted December 1992.
Accepted May 1993.

20 to 30 percent of their estimates of the coefficients for the reaction of expectations exceed their theoretical upper limit of one.

To overcome these problems, we propose an alternative method by estimating the *distributed lag* form of the model using a maximum-likelihood-based approach (Maddala 1988, 343–44; Kmenta 1986, 532–33). Using this procedure, we obtain results that are inconsistent with those of Lee and Wu (1988).

Key Words: *Forecasts, Financial ratios, Adaptive expectations, Distributed lags.*

Data Availability: *The data used in this study are available from the authors upon request.*

I. Problems with Lee and Wu's Procedure

IN this section, we introduce Lee and Wu's (1988) model and discuss the problems with their estimation procedure. Lee and Wu estimate the structural parameters of the autoregressive form of their generalized partial adjustment-adaptive expectations model:[1]

$$y_t = a_0 + a_1 x_t + a_2 y_{t-1} + a_3 y_{t-2} + \mu_t, \tag{1}$$

where:

$a_1 = \lambda\delta;$
$a_2 = 2 - \lambda - \delta;$
$a_3 = -(1-\lambda)(1-\delta);$

and $y_i(i=t,\ t-1,$ and $t-2)$ represents the current and lagged values of a firm's financial ratio; x_t represents the industry average; λ is the coefficient of the speed of adjustment; δ is the coefficient of the reaction of expectations; and a_0 and μ_t are the intercept and disturbance terms, respectively. Theoretically, λ and δ should lie between zero and one.

Lee and Wu (1988) obtain estimates of the parameters in equation (1) using the Gauss-Newton nonlinear least squares method (NLS). One problem with this approach is that the parameters λ and δ in the equation occur symmetrically, thereby producing ambiguity in their estimates (Maddala 1988, 353). Because λ cannot be distinguished from δ, neither can be estimated unambiguously. For example, if the sum of squared errors (SSE) is minimized at the estimates of $\lambda=0.5$ and $\delta=0.9$, the estimates of $\lambda=0.9$ and $\delta=0.5$ also minimize SSE. If one set of estimates is characterized by a high value of λ and a low value of δ, the other set is characterized by a low value of λ and a high value of δ. A single set of estimates occurs only if the estimates of λ and δ are equal.

The application of NLS to equation (1) results in convergence to either the single global minimum or one of the two global minima. When dual minima exist, the initial estimates used in the estimation procedure determine to which one of the minima the procedure converges. The final estimates of λ and δ are sensitive to the initial values used in their estimation.

---

[1] There are two typographical errors in Lee and Wu's (1988) equation (8): "x_1" should be "x_t" and "$-a_2y_{t-1}$" should be "$+a_2y_{t-1}$."

We were unable to obtain data from Lee and Wu (1988) and do not know which version (or versions) of the COMPUSTAT file they used in collecting their data. Attempts were made to replicate Lee and Wu's sample as closely as possible. For this comment, data for the same period used by Lee and Wu (1965–1984) was obtained from the 1984 and 1986 versions of the file. As a consequence, the data employed in this comment may differ slightly from those used by Lee and Wu. However, any differences do not appear to affect the comparison significantly.[2]

Twenty-one industries meet Lee and Wu's (1988) sample selection criteria, as compared to eight in their study. Six of the eight industries examined by Lee and Wu are readily identifiable: paper and allied products, chemicals and allied products, petroleum refining, electronic components, certified air transportation, and grocery stores. Only these six industries are used in our comment because the other two industries in Lee and Wu's study (textiles, minerals and mining) could not be identified. These six industries contain 102 firms and we examined the same six financial ratios as Lee and Wu did. The arithmetic means of the firms' financial ratios in each industry group are used as proxies for the target ratios.

The following procedure shows the effect of different combinations of initial values on the final parameter estimates. The parameter estimates for the current ratios of the 102 sample firms are obtained using Gauss-Newton NLS under 11 different sets of initial values:[3]

A. The mean parameter estimates reported by Lee and Wu (1988, 299) for the current ratio:

(1) a_0:0.0239 λ:0.7846 δ:0.8523.

B. The two sets of parameter estimates that minimize the residual sum of squares in a grid search over the following values of the parameters λ (0 to 1 by 0.1) and δ (0 to 1 by 0.1):

(2) high λ and low δ, and
(3) low λ and high δ.

C. Nine arbitrary sets of initial values:

	λ	δ
(4)	0	0
(5)	0.2	0.2
(6)	0.4	0.4
(7)	0.5	0.5
(8)	0.6	0.6
(9)	0.7	0.7
(10)	0.8	0.8
(11)	1	1

---

[2] Lee and Wu also estimated $a_0, a_1, a_2,$ and a_3 in their equation (8) using ordinary least squares (OLS). To check the effect of the possible data difference, we estimated the model using OLS. The resulting parameter estimates and standard R^2 values are not significantly different from those obtained by Lee and Wu's using OLS.

[3] The results for the other five ratios are quite similar to those for the current ratio, and are available from the authors.

The results from the above procedure indicate that for 68 firms two sets of estimates exist to which convergence is possible for the coefficients of partial adjustment (λ) and adaptive expectations (δ).[4] These sets of estimates, having identical SSE, represent dual global minima. The minimum to which the procedure converges depends on the intitial values used. For the remaining 34 firms, the estimation procedure converges to a single set of estimates for λ and δ (the global minimum) in which λ and δ are equal. In these cases, the Jacobian matrix (the matrix of partial derivatives) is singular, implying possible overparameterization of the model.

Many estimates of λ and δ for the individual firms exceed their theoretical upper limit of one. The estimates of λ and δ range from 0.04 to 1.52. The values of λ or δ exceed one for 58 firms (57 percent). This problem was also encountered by Lee and Wu (1988); approximately 25 percent of their estimates of λ and about 30 percent of their estimates of δ for the current ratio exceed one.[5]

Table 1 summarizes the mean NLS parameter estimates of the generalized model (1) for the current ratios of the 102 sample firms under the 11 different sets of initial values outlined above. The results in this table suggest two problems. First, the mean parameter estimates of λ and δ are sensitive to the initial values used in their estimation. The mean sets of estimates range from 0.49 for λ and 1.04 for δ to 1.05 for λ and 0.48 for δ. The mean SSE of 2.5814 is identical across all sets of initial values. High initial values of λ and δ result in final estimates close to those reported by Lee and Wu (1988). They appear to use a unique set of initial values, which they claim is based on Zellner and Geisel's (1970) maximum likelihood method. In the case of dual minima, the initial values Lee and Wu use favor one set of estimates for λ and δ over the other set. As a result, their mean parameter estimates for each ratio depend on the initial values used in the estimation for each firm.

Second, the coefficients of determination (R^2) for NLS reported by Lee and Wu (1988) do not appear to be the standard or conventionally calculated values. Coefficients of determination can be calculated in two ways: one for models with intercept terms (R^2), and the other for models without intercept terms (R_0^2):[6]

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum(\hat{y}_t - y_t)^2}{\sum(y_t - \bar{y})^2}, \text{ and}$$

$$R_0^2 = 1 - \frac{SSE}{\sum y_t^2} = 1 - \frac{\sum(\hat{y}_t - y_t)^2}{\sum y_t^2}.$$

The coefficients of determination from both methods for the current ratio of our sample are reported in table 1. Across firms, the mean of the standard R^2 is 0.4075, while the mean of the R_0^2 is 0.9778.

It appears that Lee and Wu (1988) reported the R_0^2 values when they applied NLS to their generalized model, but reported the R^2 values when they applied the ordinary least squares method (OLS) to this model. In fact, the standard R^2 of Lee and Wu's generalized model estimated by NLS cannot be greater than the standard R^2 of the

---

[4] The results for the individual firms could not be fully described, but are available upon request.

[5] Zellner and Geisel (1970) note that their estimate of k (which is similar to Lee and Wu's partial adjustment coefficient λ) that is larger than one will be deemed by most to be unreasonable on *a priori* grounds and may be due to the inadequacies of the model and/or the data.

[6] See Stewart (1976, 70–75) for an overview of the rationale underlying R_0^2.

Table 1
Mean Nonlinear Least Squares Estimates of the Autoregressive Form of the Generalized Model for the Current Ratio Under Different Sets of Starting Values

$$y_t = a_0 + \lambda\delta x_t + (2-\lambda-\delta)y_{t-1} - (1-\lambda)(1-\delta)y_{t-2} + \mu_t$$

(Sample Period: 1965–1984, $n = 102$)

	a_0		λ		δ	
Starting Values	*Estimate*	*t-statistic*	*Estimate*	*t-statistic*	*Estimate*	*t-statistic*
(1) *Lee and Wu's Mean Estimates*						
Mean	−0.0166	−1.2072	0.4923	1.5231	1.0389	3.9647
Std. deviation	0.2746	1.5069	0.2703	0.7447	0.2833	1.5784
(2) *Grid Search (high λ, low δ)*						
Mean	−0.0166	−1.1147	0.8816	3.3457	0.6495	2.1623
Std. deviation	0.2746	1.3663	0.3797	1.8391	0.3645	1.6319
(3) *Grid Search (high δ, low λ)*						
Mean	−0.0166	−1.1673	0.6488	2.2186	0.8824	3.2842
Std. deviation	0.2746	1.4473	0.3644	1.6159	0.3793	1.8423
(4) λ *and* δ: 0						
Mean	−0.0166	−1.2058	1.0474	4.0070	0.4838	1.4713
Std. deviation	0.2746	1.4276	0.2754	1.5385	0.2607	0.7509
(5) λ *and* δ: 0.2						
Mean	−0.0166	−1.2427	1.0474	4.0560	0.4838	1.4797
Std. deviation	0.2746	1.4077	0.2754	1.5659	0.2607	0.7058
(6) λ *and* δ: 0.4						
Mean	−0.0166	−1.1126	1.0474	3.9841	0.4838	1.4857
Std. deviation	0.2746	1.4437	0.2755	1.6398	0.2607	0.7279
(7) λ *and* δ: 0.5						
Mean	−0.0166	−1.2066	1.0177	3.9159	0.5135	1.5869
Std. deviation	0.2746	1.4601	0.3270	1.7537	0.2624	0.7980
(8) λ *and* δ: 0.6						
Mean	−0.0166	−1.1759	0.8687	3.1398	0.6625	2.3983
Std. deviation	0.2746	1.4634	0.4348	2.0126	0.3061	1.5238
(9) λ *and* δ: 0.7						
Mean	−0.0166	−1.3042	0.7776	2.8473	0.7536	2.7981
Std. deviation	0.2746	1.4725	0.4592	2.0586	0.3051	1.3527
(10) λ *and* δ: 0.8						
Mean	−0.0166	−1.2163	0.6260	2.1612	0.9052	3.3081
Std. deviation	0.2746	1.5015	0.4176	1.6944	0.3008	1.6663
(11) λ *and* δ: 1						
Mean	−0.0166	−1.1412	0.4874	1.5060	1.0438	4.1690
Std. deviation	0.2746	1.4378	0.2702	0.7733	0.2737	1.6373

Definition: $R^2 = 1 - [\Sigma(\hat{y}_t - y_t)^2 / \Sigma(y_t - \bar{y})^2]$, and $R_0^2 = 1 - [\Sigma(\hat{y}_t - y_t)^2 / \Sigma y_t^2]$.

Note: The different sets of starting values result in identical R^2, R_0^2, and SSE: 0.4075, 0.9778, and 2.5814, respectively.

model estimated by OLS, because their nonlinear model is a specific case of their more general linear model.[7]

Lee and Wu's (1988) presentation of the coefficients of determination for their model is misleading, because they estimated these coefficients by using different methods for OLS and NLS without disclosing that different methods were used in their computation. The presentation of the very high R_0^2 values for their nonlinear model and the considerably lower R^2 values for their linear model gives the impression that their nonlinear model has better explanatory power. In fact, the opposite is the case.[8]

II. An Alternative Procedure for Estimating the Parameters

Lee and Wu's (1988) generalized model, or more specifically the adaptive expectations component of the model, can be expressed in its distributed lag form. Unambiguous estimates of λ and δ can then be obtained from the application of a maximum-likelihood-based technique to this form (Maddala 1988; Kmenta 1986).

Lee and Wu's (1988) equation (6) is the distributed lag form of the generalized model. Lee and Wu apply the Koyck transformation to this equation to estimate it from a finite sample of data. However, the Koyck transformation need not be invoked. Lee and Wu's equation (6) can be expressed as:

$$y_t=(1-\lambda)y_{t-1}+\lambda\left[\sum_{i=0}^{\infty}\delta(1-\delta)^i x_{t-i}\right]. \tag{2}$$

The infinite series component of equation (2) can be split into the observed and unobserved past and expressed as:

$$\sum_{i=0}^{t-1}\delta(1-\delta)^i x_{t-i}+\sum_{i=t}^{\infty}\delta(1-\delta)^i x_{t-i}. \tag{3}$$

The first part is observed and we denote it by z_{1t}. By substituting j for $i-t$, the second part of expression (3) can be rewritten as:

$$(1-\delta)^t\sum_{i=t}^{\infty}\delta(1-\delta)^{i-t}x_{t-i}=(1-\delta)^t\sum_{j=0}^{\infty}\delta(1-\delta)^j x_{-j}. \tag{4}$$

Treating the summation component of equation (4) as an unknown parameter c and defining $z_{2t}=(1-\delta)^t$, we can write expression (3) as:

$$z_{1t}+cz_{2t}. \tag{5}$$

If we introduce an intercept and a disturbance term, equation (2) can now be expressed as:

$$y_t=a_0+(1-\lambda)y_{t-1}+\lambda z_{1t}+c'z_{2t}+\mu_t, \tag{6}$$

where $c'=\lambda c$, a_0 is the intercept term, and μ_t is the disturbance term. The variables z_{1t} and z_{2t} depend on δ, and therefore δ has to be estimated along with a_0, $(1-\lambda)$, λ, and c'.

---

[7] Estimates of λ and δ can be extracted from the OLS estimates of a_1 and a_2 in the generalized model (1) after imposing the following restriction on its parameters: $a_3=1-a_1-a_2$.

[8] The standard R^2 values for Lee and Wu's linear model are all greater than those for their nonlinear model.

This estimation can be achieved using an approach based on the maximum likelihood principle.

Assuming that the disturbance term in equation (6) is normal and independently distributed, the logarithmic likelihood function for $y_1, y_2, \ldots, y_n$ is:

$$L = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{n}[y_t - a_0 - (1-\lambda)y_{t-1} - \lambda z_{1t} - c'z_{2t}]^2. \tag{7}$$

Maximizing L with respect to a_0, $(1-\lambda)$, λ, and c' is equivalent to minimizing

$$\sum_{t=1}^{n}[y_t - a_0 - (1-\lambda)y_{t-1} - \lambda z_{1t} - c'z_{2t}]^2, \tag{8}$$

with respect to the same parameters. This process can be accomplished by regressing y_t on y_{t-1}, z_{1t}, and z_{2t} for various values of δ within its theoretical range $(0,1)$. At these chosen values of δ, the following variables are calculated:

$$z_{1t} = \sum_{i=0}^{t-1}\delta(1-\delta)^i x_{t-i}, \text{ and}$$

$$z_{2t} = (1-\delta)^t.$$

Thus,

$$\begin{aligned} z_{11} &= \delta x_1, & z_{21} &= (1-\delta)^1, \\ z_{12} &= \delta[x_2 + (1-\delta)x_1], & z_{22} &= (1-\delta)^2, \\ z_{13} &= \delta[x_3 + (1-\delta)x_2 + (1-\delta)^2 x_1], & z_{23} &= (1-\delta)^3, \end{aligned}$$

and so on.

We estimate equation (6) using OLS over the range $(0,1)$ for values of δ at intervals of 0.05 and obtain the residual sum of squares $[RSS(\delta)]$. The value of δ for which the minimum $RSS(\delta)$ is obtained serves as the starting point for the next stage, where the same process is repeated over a range of 0,1 at intervals of 0.01 for δ. The value of δ for which the $RSS(\delta)$ is a minimum is used as the estimate of δ, and the estimates of a_0 and λ are obtained from the least squares method. The asymptotic standard error of these parameters is estimated from the square roots of the diagonal elements of the inverse information matrix using maximum likelihood functions to value these parameters at their estimates.

Table 2 summarizes the cross-sectional distribution of the parameter estimates for the six financial ratios obtained from the distributed lag form of the generalized model.[9] Most of the estimated coefficients of partial adjustment (λ) fall into the theoretical range between zero and one with significantly fewer exceptions than Lee and Wu's (1988) estimates. The estimated coefficients of adaptive expectations (δ) all fall between zero and one because of their method of estimation. The mean R^2 values of the distributed lag form of the generalized model range from 0.51 to 0.69.

---

[9] We also examined the error structure of the distributed lag form of the model. Similar to the error structure of the autoregressive form of Lee and Wu's model, our first order autoregressive coefficients of the financial ratios are generally small and insignificant. The parameter estimates are mostly insensitive to treatment for autocorrelation. Thus, for comparison with Lee and Wu's results, we report only those without treatment for autocorrelation.

Table 2

Parameter Estimates From the Distributed Lag Form of the Generalized Model

$y_t=a_0+(1-\lambda)y_{t-1}+\lambda z_{1t}+c'z_{2t}+\mu_t$, where $z_{1t}=\Sigma\delta(1-\delta)^i x_{t-i}$, and $z_{2t}=(1-\delta)^t$.

(Sample Period: 1965–1984, $n=102$)

	a_0		(1–λ)		λ		c′		δ		
Ratio	*Est.*	*t-stat.*	*Est.*	*t-stat.*	*Est.*	*t-stat.*	*Est.*	*t-stat.*	*Est.*	*t-stat.*	R²
Quick Ratio											
Mean	–0.403	–0.108	0.299	1.321	0.701	0.854	–3.021	0.298	0.48	0.988	0.511
Std. deviation	1.906	0.340	0.235	1.221	0.235	0.891	22.268	0.708	0.44	0.976	0.222
Fractiles											
0.1	–2.598	–0.473	–0.034	–0.141	0.429	0.010	–24.636	–0.026	0.01	0.010	0.221
0.2	–0.665	–0.340	0.112	0.479	0.494	0.020	–8.596	–0.018	0.01	0.029	0.318
0.3	–0.298	–0.141	0.183	0.750	0.585	0.085	–1.321	–0.009	0.03	0.114	0.387
0.4	–0.204	–0.062	0.270	1.020	0.640	0.251	0.121	0.002	0.10	0.395	0.443
0.5	–0.100	–0.026	0.318	1.355	0.682	0.536	0.997	0.012	0.36	0.897	0.500
0.6	–0.045	–0.008	0.360	1.470	0.730	0.897	1.568	0.026	0.82	1.157	0.591
0.7	0.031	0.009	0.415	1.676	0.817	1.313	2.515	0.072	0.99	1.522*	0.661
0.8	0.139	0.063	0.506	2.025	0.888	1.745*	4.236	0.216	0.99	1.867**	0.728
0.9	0.471	0.199	0.571	2.696	1.034	2.289**	10.454	1.312	0.99	2.175**	0.791
Current Ratio											
Mean	–0.671	–0.102	0.274	1.159	0.726	0.812	–5.312	0.509	0.38	0.808	0.533
Std. deviation	2.484	0.319	0.243	1.138	0.243	1.033	35.261	0.968	0.40	0.974	0.253
Fractiles											
0.1	–4.099	–0.357	–0.076	–0.299	0.432	0.004	–26.349	–0.016	0.01	0.004	0.165
0.2	–1.055	–0.230	0.081	0.328	0.524	0.009	–0.555	–0.008	0.01	0.015	0.274
0.3	–0.594	–0.142	0.152	0.610	0.571	0.027	0.824	0.003	0.01	0.039	0.387
0.4	–0.358	–0.052	0.231	0.889	0.670	0.139	1.386	0.013	0.08	0.248	0.473
0.5	–0.201	–0.015	0.294	1.086	0.706	0.375	1.932	0.051	0.21	0.535	0.537
0.6	–0.074	–0.003	0.330	1.388	0.769	0.751	2.480	0.171	0.33	0.775	0.616
0.7	–0.006	–0.000	0.429	1.759	0.848	1.114	3.674	0.460	0.65	1.066	0.707
0.8	0.168	0.033	0.476	2.070	0.919	1.502*	5.517	0.831	0.99	1.474*	0.791
0.9	0.532	0.157	0.568	2.548	1.076	2.155**	8.845	2.097	0.99	2.149**	0.866
Equity/Total Debts											
Mean	–0.438	–0.068	0.481	2.088	0.519	0.635	–2.146	0.302	0.37	1.004	0.652
Std. deviation	3.110	0.562	0.253	1.314	0.253	0.873	33.169	0.777	0.40	1.235	0.224
Fractiles											
0.1	–1.497	–0.514	0.143	0.622	0.181	0.004	–16.895	–0.056	0.01	0.010	0.331
0.2	–0.570	–0.256	0.287	1.164	0.311	0.013	–2.556	–0.018	0.01	0.023	0.470
0.3	–0.339	–0.136	0.363	1.516	0.391	0.028	–0.494	–0.010	0.01	0.057	0.546
0.4	–0.158	–0.043	0.435	1.819	0.438	0.102	0.000	0.000	0.07	0.247	0.613
0.5	–0.102	–0.019	0.492	2.086	0.508	0.239	0.502	0.020	0.20	0.665	0.691
0.6	–0.010	–0.002	0.562	2.244	0.565	0.525	1.232	0.041	0.36	0.993	0.742
0.7	0.106	0.013	0.609	2.412	0.637	0.780	1.793	0.102	0.65	1.254	0.801
0.8	0.489	0.062	0.689	2.900	0.713	1.130	2.549	0.396	0.99	1.875**	0.860
0.9	0.955	0.117	0.819	3.546	0.857	1.773**	6.891	1.319	0.99	2.975**	0.917

Continued

Table 2—*Continued*

	a_0		$(1-\lambda)$		λ		c'		δ		
Ratio	*Est.*	*t-stat.*	*Est.*	*t-stat.*	*Est.*	*t-stat.*	*Est.*	*t-stat.*	*Est.*	*t-stat.*	R^2
Sales/Inventory											
Mean	7.154	−0.005	0.323	1.244	0.677	0.661	−3.638	0.287	0.33	0.711	0.607
Std. deviation	19.569	0.282	0.289	1.278	0.289	0.841	127.974	0.623	0.40	0.829	0.252
Fractiles											
0.1	−3.072	−0.232	−0.011	−0.044	0.311	0.009	−52.825	−0.038	0.01	0.017	0.262
0.2	−1.456	−0.082	0.071	0.298	0.406	0.016	−21.807	−0.029	0.01	0.032	0.322
0.3	−0.542	−0.017	0.143	0.552	0.543	0.028	−12.317	−0.019	0.01	0.050	0.445
0.4	0.006	0.001	0.243	0.908	0.646	0.081	−3.895	−0.011	0.01	0.147	0.591
0.5	0.609	0.016	0.313	1.254	0.687	0.237	1.391	0.006	0.14	0.318	0.681
0.6	1.440	0.031	0.354	1.460	0.757	0.492	4.463	0.018	0.22	0.648	0.740
0.7	4.898	0.048	0.457	1.714	0.857	0.734	9.895	0.155	0.35	1.109	0.808
0.8	15.704	0.075	0.594	2.012	0.929	1.411*	14.903	0.564	0.96	1.566*	0.851
0.9	25.127	0.158	0.689	2.740	1.011	2.049**	29.632	1.138	0.99	1.794**	0.893
Sales/Total Assets											
Mean	−0.045	−0.039	0.401	1.683	0.599	1.129	−1.280	0.221	0.58	1.228	0.692
Std. deviation	1.311	0.507	0.239	1.154	0.239	1.017	17.626	0.556	0.44	1.100	0.186
Fractiles											
0.1	−0.912	−0.498	0.159	0.438	0.311	0.011	−12.063	−0.023	0.01	0.026	0.428
0.2	−0.218	−0.260	0.220	0.829	0.376	0.043	−6.442	−0.016	0.01	0.071	0.538
0.3	−0.109	−0.144	0.276	1.081	0.469	0.284	−1.374	−0.010	0.15	0.517	0.615
0.4	−0.074	−0.043	0.343	1.225	0.549	0.580	−0.001	−0.000	0.26	0.932	0.671
0.5	−0.029	−0.005	0.407	1.476	0.593	0.908	0.190	0.004	0.85	1.150	0.726
0.6	0.019	0.012	0.451	1.936	0.657	1.279	0.754	0.019	0.99	1.295	0.777
0.7	0.100	0.041	0.531	2.255	0.724	1.753**	1.506	0.039	0.99	1.635*	0.849
0.8	0.171	0.142	0.624	2.572	0.780	2.187**	4.446	0.136	0.99	2.021**	0.870
0.9	0.662	0.352	0.689	2.914	0.841	2.572**	7.244	0.839	0.99	2.462**	0.894
Net Operating Income/Total Assets											
Mean	−0.030	−0.081	0.315	1.452	0.685	1.334	0.023	0.159	0.72	1.456	0.578
Std. deviation	0.133	0.601	0.241	1.278	0.241	0.946	2.878	0.555	0.41	0.946	0.201
Fractiles											
0.1	−0.130	−0.821	−0.001	−0.003	0.381	0.015	−2.755	−0.027	0.01	0.034	0.313
0.2	−0.034	−0.495	0.051	0.183	0.482	0.191	−1.672	−0.018	0.12	0.444	0.394
0.3	−0.018	−0.253	0.183	0.689	0.538	0.743	−0.658	−0.011	0.53	0.994	0.479
0.4	−0.013	−0.124	0.273	1.041	0.612	1.079	−0.008	−0.007	0.99	1.211	0.559
0.5	−0.004	−0.038	0.336	1.321	0.664	1.244	0.008	0.001	0.99	1.550*	0.626
0.6	0.001	0.013	0.388	1.732	0.727	1.589*	0.194	0.017	0.99	1.673*	0.659
0.7	0.011	0.117	0.462	1.923	0.817	1.922**	0.466	0.026	0.99	2.076**	0.718
0.8	0.017	0.287	0.518	2.435	0.949	2.167**	1.135	0.037	0.99	2.345**	0.769
0.9	0.041	0.544	0.619	3.342	1.001	2.630**	3.627	0.435	0.99	2.650**	0.795

Definition: $R^2=1-[\Sigma(\hat{y}_t-y_t)^2/\Sigma(y_t-\bar{y})^2]$.

* Significant at the 0.10 level.

** Significant at the 0.05 level.

The proportions of significant coefficients of partial adjustment and adaptive expectations are generally slightly less than those reported by Lee and Wu (1988). The partial adjustment coefficient (λ) are significant (at the five percent level) for about 30 percent of the firms in the sample for both the net operating income/total assets and sales/total assets ratios. For the remaining financial ratios, the partial adjustment coefficients are generally not significant at the five percent level.

The coefficients of adaptive expectations (δ) are significant (at the five percent level) for about 30 percent of the sample firms for the net operating income/total assets ratio, and 20 percent for the sales/total assets, equity/debt, and quick ratios. The coefficients for the remaining financial ratios are generally insignificant.

The averages of λ for the six ratios range from 0.52 to 0.73, all lower than those of Lee and Wu (1988) which range from 0.68 to 0.81. The biggest difference is in the mean estimate for the equity/total debts ratio: our 0.52 versus Lee and Wu's 0.81. The average values for δ range from 0.33 to 0.72. They are also all lower than those of Lee and Wu which range from 0.78 to 0.85. Moreover, in contrast to Lee and Wu's estimates, more than half of our estimates of δ for each ratio are close to either zero or one, suggesting that the majority of firms either (1) do not periodically adjust their target levels once such targets are set, or (2) adjust target ratios to their industry averages instantly (in this case, finanical ratio adjustment follows a simple partial adjustment process).

However, the coefficients of partial adjustment and adaptive expectations need to be interpreted with caution. Before Lee and Wu (1988) applied NLS to estimate λ and δ in their equation (8), they estimated a_0, a_1, a_2, and a_3 in the equation by OLS. Their purpose was to determine whether the coefficient of y_{t-2} (a_3) indicated the presence of an expectation adjustment lag. However, most of the coefficients of y_{t-2} were insignificant. Lee and Wu (p. 296) finesse this issue by stating: "The variable of the financial ratio lagged for two periods, y_{t-2}, also plays a role in the ratio adjustment process." The lack of significance of a_3 brings into question the theoretical foundations of Lee and Wu's multi-year model and suggests that a simple model like the following may be more appropriate to characterize the ratio adjustment process:

$$y_t = b_0 + b_1 x_t + b_2 y_{t-1} + \mu_t, \tag{9}$$

where y_t, x_t, and y_{t-1} are defined in equation (1), and b_i ($i=0$, 1, and 2) represents regression coefficients. This model causes a serious problem for those who claim that the partial adjustment-adaptive expectations model proves that firms adjust their ratios to industry averages. In equation (9), b_1 is an industrywide effect common to all firms. Therefore, it is impossible to distinguish between industrywide effects and firms' adjustments to industry averages.

III. Conclusions

It is inappropriate for Lee and Wu (1988) to estimate the structural parameters of the autoregressive form of their partial adjustment-adaptive expectations model by using the Gauss-Newton nonlinear least squares method. Specifically, this approach leads to ambiguity in the estimate of the parameters: the speed of adjustment and the reaction of expectations. The parameter estimates obtained from such a method depend directly on the initial values used in the estimation. In addition, many of Lee and Wu's (1988) estimates of the parameters exceed their theoretical upper limit of one.

Using the maximum likelihood method to estimate the parameters in the distributed lag form of the generalized model corrects these problems and provides some new findings. First, unlike Lee and Wu's (1988) results where most coefficients of adaptive expectations are roughly evenly distributed between zero and one, most of our estimates of the coefficients are close to either zero or one. Second, our average estimates of the speed of the adjustment and the reaction of expectations for each ratio are smaller than those of Lee and Wu. Finally, our significance levels of partial adjustment and adaptive expectations for each ratio are generally slightly less than those of Lee and Wu.

References

Kmenta, J. 1986. *Elements of Econometrics*. 2d ed. New York: Macmillan.

Lee, C. F., and C. Wu. 1988. Expectation formation and financial ratio adjustment processes. *The Accounting Review* 63 (April): 292–306.

Lev, B. 1969. Industry averages as targets for financial ratios. *Journal of Accounting Research* 7 (Autumn): 290–99.

Maddala, G. 1988. *Introduction to Econometrics*. New York: Macmillan.

Stewart, J. 1976. *Understanding Econometrics*. London: Hutchison.

Zellner, A., and M. S. Geisel. 1970. Analysis of distributed lag models with applications to consumption function estimation. *Econometrica* 38 (November): 865–88.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 953–955

Expectation Formation and Financial Ratio Adjustment Processes: A Reply

Cheng F. Lee
Rutgers University
Chunchi Wu
Syracuse University

TUNG and Crowe (1993) in their comment on our article (1988) suggest an alternative method based on the distributed lag form of adaptive expectations for estimating the parameters of the partial adjustment and adoptive expectations model. Tung and Crowe claim that the method has resolved the ambiguity problem of parameter estimation. The results in the comment, however, show that all of the average *t*-statistics of the partial adjustment coefficient (λ) and adaptive expectations coefficient (δ) for six financial ratios are insignificant. Therefore, financial ratio adjustment does not exist and the findings of Lev (1969) and Frecka and Lee (1983) are statistical artifacts at best.

In response, we indicate pitfalls of the proposed statistical procedure and suggest a possible extension. We also reply to some of the unfounded concerns raised by Tung and Crowe (1993).

I. Potential Problems of the Alternative Estimation Procedure

There are several concerns about the alternative estimation method. The estimation of parameters is based on the following equation:

$$\begin{aligned} y_t &= a_0+(1-\lambda)y_{t-1}+\lambda\delta\sum_{i=0}^{t-1}(1-\delta)^i x_{t-i}+\lambda\delta(1-\delta)^t\sum_{j=0}^{\infty}(1-\delta)^j x_{-j}+\mu_t, \\ &= a_0+(1-\lambda)y_{t-1}+bz_{1t}'+bc+z_{2t}+\mu_t, \\ &= a_0+(1-\lambda)y_{t-1}+\lambda z_{1t}+c'z_{2t}+\mu_t, \end{aligned} \tag{1}$$

Invited by the Editor.
Received May 1993.

where $b=\lambda\delta$; $z_{2t}=(1-\delta)^t$; $z_{1t}=\delta z'_{1t}=\delta\Sigma_i(1-\delta)^i x_{t-i}$; $c'=\lambda\delta c^*=\lambda\delta\Sigma_j(1-\delta)^j x_{-j}$; y_t and x_t are the firm's financial ratio and industry average; a_0 is a constant; and μ_t is the disturbance term. The last expression in equation (1) is identical to equation (6) of the comment, while the second expression is used to highlight the nonlinearity in λ and δ. In the above equation, there are still two parameters (λ and δ) and three restrictions, and the estimation problem remains. First of all, for example, $c'=\lambda\delta c^*$. Thus, a higher λ and a lower δ can still result in the same sum of squared errors (SSE) as is the case with a lower λ and a higher δ. Second, the estimates of parameters by the suggested iterative ordinary least squares (OLS) method will be asymptotically efficient only if we have no information about x other than the sample values of $x_1, x_2, \ldots, x_n$ (Kmenta 1986, 533). However, there are only 20 yearly observations in the annual COMPUSTAT tape. Other earlier information is available but is not included in the tape. The relatively small sample size restricted by an annual COMPUSTAT tape can cause estimation inefficiency and increase standard errors of parameter estimates. Third, the OLS estimator is inconsistent if the disturbance term μ_t is assumed to be serially uncorrelated when it is actually serially correlated. This problem is particularly cumbersome for regressions including many firms. If the distrubance term is not independent for some firms and this serial correlation problem is not handled properly, the average parameter estimates computed from the entire sample will be misleading. In contrast, for the nonlinear regressions in our original paper, the normality of the disturbance term is not crucial (Malinvaud 1975, 325–41).

The above problems may have caused the poor performance of Tung and Crowe's (1993) model. For example, c' is supposed to be positive for all financial ratios with the possible exception of the ratio of net operating income to total assets. Yet, table 2 shows that c' estimates are negative for a high percentage of firms. For instance, for the sales/inventory and sales/total asset ratios, more than 40 percent of c' estimates are negative. Furthermore, the partial adjustment coefficient (λ) estimates are mostly insignificant, which contrasts sharply with the results in Lev (1969) and Frecka and Lee (1983). The lower t-statistics are likely caused by the statistical problems mentioned above. The correct asymptotic standard errors for λ and δ for a regression model similar to that suggested by Tung and Crowe (1993) were derived by Doran and Griffiths (1978). It is unclear whether Tung and Crowe have calculated the standard errors using the asymptotic formula.

Furthermore, some of the results in table 1 of the comment appear to contradict intuition. Table 1 indicates that the values of R^2, sum of squared errors (SSE), and the constant term a_0 are identical for all regressions from (1) to (11). A sufficient condition for equal SSE values is that a_1, a_2, and a_3 are identical across regressions in table 1, where $a_1=\lambda\delta$, $a_2=(2\lambda-\delta)$, and $a_3=(1-\lambda)(1-\delta)$, since a_0 is identical. However, the values of a_1, a_2, and a_3 in table 1 do not satisfy this condition. For instance, $a_1=0.5114$, $a_2=0.4688$, and $a_3=0.0197$ in the first set of regressions, and $a_1=0.5725$, $a_2=0.4689$, and $a_3=0.0414$ in the second set. Thus, the results for SSE and R^2 in table 1 appear to be counterintuitive.

II. Explanations of Previous Results

Our data were drawn from the annual industrial file of the COMPUSTAT tape, and therefore can be easily retrieved. Tung and Crowe (1993) indicate that our OLS results for parameter estimates and R^2 are very similar to their replications (Tung and Crowe

1993, fn. 2). Our nonlinear estimation was conducted using the SAS subroutines, and the R^2 values were obtained directly from the SAS output file. Tung and Crowe need not have the strange thought that we intend to mislead readers by adopting different R^2 definitions. Also, in our OLS regressions, the *t*-statistics of y_{t-2} are often greater than one in absolute values for about 40 percent of the sample firms. A variable will add significantly to the explanatory power when its *t*-value in absolute terms is greater than one. Thus in practice, including y_{t-2} in the regression model improves future forecasts. This usage is precisely what we meant by the "role" of y_{t-2}, which clearly has nothing to do with finessing as Tung and Crowe assert. They are incorrect in arguing that expectations are not formed adaptively based on the OLS coefficient estimates and *t*-statistics of y_{t-2} (Doran and Griffiths 1978, 142).

III. Summary and Extension

The method suggested by Tung and Crowe (1993) represents one way to handle the problem of symmetry when the error term is well behaved and the sample size is sufficiently large. Other alternative methods are also available. As indicated by Maddala (1988), the ambiguity (symmetry) problem arises only if the model is estimated by ordinary least squares while assuming that the error terms are serially uncorrelated. Our nonlinear regression method, accompanied by proper initial values, was intended to avoid the problem confronted by the OLS estimation (Waud 1968). The ambiguity problem will disappear if the autocorrelation structure of the error term (equal to $u_t-[1-\delta]u_{t-1}$ for our eq. [8]) is accounted for in the nonlinear regression. As mentioned in our previous article and indicated in the comment, the nonlinear regression results with and without autocorrelation appear to be quite close. This outcome suggests that the nonlinear regression results are reasonably reliable. Alternatively, including some other explanatory variables in the model can also resolve the ambiguity problem (Waud 1968; Maddala 1988, 353–54). In the present case, potential candidates could be firm size or other firm characteristics which are likely to affect financial ratio choices. Inclusion of additional explanatory variables has the advantage of resolving the symmetry problem while preserving the parsimonious feature of the partial adjustment-adaptive expectations model.

References

Doran, H. E., and W. E. Griffiths. 1978. Inconsistency of the OLS estimator of the partial adjustment-adaptive expectations model. *Journal of Econometrics* 26 (issue): 133–46.

Frecka, T. J., and C. F. Lee. 1983. Generalized financial ratio adjustment processes and their implications. *Journal of Accounting Research* 21 (Spring): 308–16.

Kmenta, J. 1986. *Elements of Econometrics*. 2d ed. New York: Macmillan.

Lee, C. F., and C. Wu. 1988. Expectation formation and financial ratio adjustment processes. *The Accounting Review* 63 (April): 292–306.

Lev, B. 1969. Industry average as targets for financial ratios. *Journal of Accounting Research* 7 (Autumn): 290–99.

Maddala, G. S. 1988. *Introduction to Econometrics*. New York: Macmillan.

Malinvaud, E. 1975. *Statistical Methods of Econometrics*. Amsterdam: North Holland.

Tung, S., and J. Crowe. 1993. Expectation formation and financial ratio adjustment processes: A comment and an extension. *The Accounting Review* 68 (October): 942–52.

Waud, R. N. 1968. Misspecification in the "partial adjustment" and "adaptive expectations" models. *International Economic Review* 9 (June): 204–17.

Zellner, A., and M. S. Geisel. 1970. Analysis of distributed lag models with applications to consumption function estimation. *Econometrica* 38 (November): 865–88.

THE ACCOUNTING REVIEW
Vol. 68, No. 4
October 1993
pp. 956-967

Book Reviews

Shane R. Moriarity, Editor

Editor's Note: Two copies of books for review should be sent to Dr. James Boatsman, Office of the Chief Accountant, U.S. Securities and Exchange Commission, Washington, D.C. 20549. The policy of the *Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

RAJ AGGARWAL, editor, *Capital Budgeting Under Uncertainty* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1993, pp. x, 292, $53.33).

This book is a collection of readings identifying problems in evaluating capital projects and suggesting methods for improving the accuracy of the capital budgeting decision. Each reading points out significant real world conditions which impact project value and which are not adequately addressed in the traditional capital budgeting model. Feasible methods are then suggested for valuing projects affected by these conditions.

Raj Aggarwal has put together a collection of 15 readings representing important valuation issues for capital budgeting in today's business environment. Several articles address strategic implications of projects, in terms of flexibility under changing economic conditions, timing and sequencing of projects within an investment program, synergy with other projects, and intangible benefits. Of particular interest to management accountants today is the article on evaluation of flexible manufacturing systems as part of a company's overall strategic plan. Other articles look at measurement problems which arise under conditions of incomplete information in such areas as utility preferences and probability distributions. The impacts of alternative competitive environments, serially dependent cash flows, replacement with unequal lives, and the learning curve phenomenon are also covered. Each reading discusses the valuation problems involved, and then suggests an approach to more accurately evaluate capital projects under these conditions. Most also illustrate the suggested approach, either with an artificial example or with real data. References listed at the end of each article invite further inquiry into areas of particular interest.

The readings were written expressly for this book, and the majority of the authors are finance academics. Most hold positions at U.S. institutions, although Israel, Spain, Finland, and Canada are also represented. Aggarwal provides an introductory chapter summarizing the state of the capital budgeting literature. This article alone makes the book worthwhile to anyone looking for a readable, nonmathematical analysis of the field of capital budgeting, including issues relating to inflation, multidivisional and multinational companies, industry structure, agency relationships, and market imperfections.

This collection of readings provides implementable procedures for evaluating capital projects under real conditions often ignored using traditional valuation methods. Although the typical corporate finance text may discuss many of these issues, usually little is offered in terms of specific procedures on how to modify the capital budgeting process accordingly. The articles in this collection emphasize the feasibility of the suggested procedures and their value in terms of improving the capital budgeting decision. In the preface, Aggarwal describes the book as suitable for advanced undergraduate and graduate courses in corporate and managerial finance or strategic planning, and for managers involved in the capital budgeting process. In addition, it would be a useful supplementary text in any advanced management accounting course which addresses the capital budgeting process.

SUSAN S. HAMLEN
Associate Professor of Accounting
State University of New York at Buffalo

DAVID GRAYSON ALLEN and KATHLEEN McDERMOTT, *Accounting Success: A History of Price Waterhouse in America 1890–1990* (Boston: Harvard Business School Press, 1993, pp. xx, 373, $35.00).

The history of Price Waterhouse & Company (PW) in the United States is presented from its beginnings as an Agency of the London, England, office of Price Waterhouse to its current status as a preeminent United States firm. PW is now a major component of "PW World Firm," one of the smallest of the large, worldwide accounting organizations. Eight chapters, organized chronologically, recount internal policy and personnel decisions as well as external legal and economic influences on the firm.

The Agency, established in New York in 1890, evolved into an independent American entity. By 1901, the London office held only a one-third interest in PW. In the early 1930s, the British financial interest ended. In the early days of Agency status, the work done by the agency resulted in reports signed with the PW name only when the British firm authorized such signing. Initially, British-trained accountants staffed the U.S. offices because London feared that their clients would object if the accounts were examined by Americans.

Three periods in the life of PW in the United States are identified: (1) 1890–1926, when the Agency began operations through the Americanization of the practice; (2) 1926–1969, when PW grew and assumed major professional leadership roles; and (3) 1969–date, when PW responded to increased competition, globalization of business, major legal challenges, and changes in the set of services provided to clients. Events in each period, such as admission of partners, changes in leadership, the contributions of Firm members to professional thought, and the Firm's adaptations to the changing economic environment, are recounted in depth.

Key firm operating philosophies and their changing/unchanging status are reported. These included: (1) the autonomy of the individual partner and offices in a decentralized organization—later changed to more centralized organization with controls over specified local office operations and decisions; (2) generalist individual partners who assumed responsibility for all types of engagements—later changed to a firm in which specialization was important; and (3) insistence on building and maintaining a reputation based on integrity and quality of service—never changed.

Contributions to the growth of the profession by leading partners are recognized. For example, the very important role played by George O. May (1902) in leading the profession toward substantially increased disclosures accompanying financial statements, shaping regulatory reforms in the 1930s, developing accounting principles (e.g., work with the AIA's Committee on Accounting Procedure), and leading development of formalized audit standards is highlighted. Other PW partners reported as making substantial contributions to the profession and the firm included Arthur Lowes Dickinson (1901), John B. Inglis (1939), Paul Grady (1944), Herman Bevis (1946), John Biegler (1958), and Joseph E. Connor (1967). Date of admission to partnership is noted in parentheses.

Major influences on PW and the accounting and auditing profession are presented, including: (1) advent of Federal income taxation, (2) Federal securities legislation, (3) McKesson-Robbins proceedings, (4) information technology development, (5) mergers and restructuring of business, (6) rise of the litigious environment, and (7) globalization of business. Effects on the profession are noted with the specific response of PW being identified. The ability of PW to achieve and retain its leading position as auditor for more New York Stock Exchange listed companies over the years than any other firm is duly noted.

This well-written, well-researched history provides a wealth of information about Price Waterhouse & Company and the development of public accounting in the United States. Extensive footnoting enables the reader to seek source information on many events and issues. Implications for the future are not developed in depth. Human interest aspects are generally excluded.

R. GLEN BERRYMAN
Professor of Accounting
University of Minnesota

KWABENA ANYANE-NTOW, editor, *International Handbook of Accounting Education and Certification* (Oxford, UK: Pergamon Ltd., 1992, pp. xxxii, 556, £65.00).

The preface to this book declares the book's intention is "to describe the educational and professional accreditation procedures in a representative sample of countries from various parts of the world" (p. xxi). The idea for this book was fostered by perceived substantial differences in the educational processes of the various nations and in the relationship between teaching institutions and professional bodies.

The book's editor is to be commended upon the clear instructions he gave to his authors, including requests to cover educational offerings at various levels, professional examinations, experience requirements, and continuing professional education; he also asked contributors to make a critical assessment of accounting education in light of socioeconomic needs. In general, the contributors followed instructions, thereby covering the subject matter in a fairly uniform manner. Western Europe is particularly well represented in this volume, along with Asia and Africa. Unfortunately, coverage of Latin America and of the Arab states is skimpy: Guyana and Peru represent the lone South American entries.

The content of the book has been leavened by several well-written contributions that reflect seasoned insights. For example, Gary Sundem has given an informative overview of the shift in U.S. accounting education toward greater emphasis on problem solving, communication skills, awareness of the business environment, and life-long learning. Adolf Enthoven displayed sensitivity toward accounting history in his erudite description of accounting and auditing in the Peoples Republic of China. Shuaib A. Shuaib pointed out that "the environmental factors surrounding business activities in developing nations differ from those in developed nations (and) that the legal, political, social, cultural, and economic systems differ from one environment to another" (p. 122).

The vital role played by accountants in the formulation, implementation, and evaluation of national development projects was made clear by several contributors. Some pointed out that accounting professionals and educators should and must address changing needs. The themes of quality, ethics, and internationalism appear to have achieved universal penetration. Integration of accounting studies could represent an emerging trend: Jake Claret stressed the integration of accounting subjects into the "real world environment" (p. 460); Andre Zund observed that Switzerland's University of St. Gall prepares students equally for both internal and external auditing; the reviewer believes that program modules covering different professional niches could enhance students' skills, to say nothing of employment opportunities.

In accordance with the charge given them, most of the contributors observed reportorial approaches; not surprisingly, criticism tended to be sparse. Few mentioned the timely question: to whom do auditors owe a duty of care; the issue whether accountants should take a leadership role in controlling bureaucratic corruption was broached by but a few; and, with the exception of some excellent comments pertaining to the Philippines, little attention was directed toward training for today's technological environment, or toward the use of technology by teachers.

I liked this book for the clear, crisp quality of its writing, and its assemblage of information hitherto not readily available. The book should be of interest to academics concerned with joint programs with foreign universities; to persons interested in teaching, research, or post-doctoral work abroad; to practitioners seeking to enter practice in particular nations; and to financial statement users or managers who recruit foreign-trained accountants. The book also presents timely information on European accountants and on differences in their training; this compilation should prove valuable as the European Community concept comes to fruition.

FELIX POMERANZ
Lecturer
Florida International University

ROBIN COOPER, ROBERT S. KAPLAN, LAWRENCE S. MAISEL, EILEEN MORRISSEY, and RONALD N. OEHN, *Implementing Activity-Based Cost Management: Moving from Analysis to Action* (Montvale, NJ: Institute of Management Accountants, 1992, pp. xviii, 336, $35.00).

Those searching for a definitive series of case studies designed to describe in detail ABC system implementations will not find them in this book. What will be found is a limited discussion concerning eight case study sites and related experiences of firms implementing "Activity-Based Cost Management Systems."

The case sites included five manufacturing organizations, a financial services company, the finance department of a large energy company, and a distribution company. None of the sites represented a complete ABC model across all expense categories. However, the authors state that "in aggregate, the eight sites... provide insight on how an enterprise-wide, organizational-activity-based model can be implemented, and on the nature of the decisions that would be supported by such a model" (p. 2).

All of the sites appeared to be KPMG Peat Marwick clients and three of the authors were, or are, associated with KPMG Peat Marwick. Futhermore, the ABC software used in each of the eight ABC implementations appeared to be the same, or very similar. The discussion usually referred to Robin Cooper's cost hierarchy of unit, batch, product sustaining, and facility sustaining activities. Thus, an advantage of the study is that the ABC system used at each of the sites appeared to be very consistent. On the other hand, the apparent close association of the researchers/authors with the case study companies, the ABC software, and the ABC implementations could make one wonder about their independence.

The ABC implementations were "off the books" in the sense that the implementations took place without disturbing the "traditional" system. "No modification to existing financial systems was required, and companies continued to run all their existing systems in parallel with their new ABC models.... *The activity-based model was treated as a management information system, not as part of the accounting system*" (p. 7).

Although the consistency of the ABC model implemented across sites was advantageous, the "traditional" systems to which the ABC model was compared were not fully explained. Thus, the reader has to view the study's conclusions with some caution since it appears a consistent ABC model was compared to eight different traditional systems. In only one of the descriptions of a traditional system (Kraft USA) was there enough information to identify the system. It was a direct costing/profit contribution type of system. Without additional detail on the traditional systems employed, the reader could legitimately wonder whether a series of minor adjustments or a few major adjustments to each traditional system would be as appropriate as a move to ABC.

Furthermore, the ABC model in this study was not explained in detail. It appears to be a refined absorption costing model which leads to a "total average cost per unit" for all products or other cost objects. However, the authors never address the issue of which volume base (practical capacity, normal volume, expected activity, or actual activity) to use in unitizing costs of batch, product sustaining, and facility sustaining activities. The authors also never address whether or not a two or more part transfer price is used in calculating the "total average cost per unit" which could indicate that the ABC is a variation of the direct costing/profit contribution ideology.

The reviewer wishes that fewer sites were involved—perhaps a maximum of two or three. In such a case, more detail could have been added in the body of the study or in appendices. Such detail could have resolved the above questions.

To the reviewer, some of the more interesting findings of these case studies were as follows:

1. "Information on the cost and profitability of products and customers yield the usual ABC insights: a few products and customers accounted for the great bulk of profits..." (p. 305). "Consistent with the fundamentals of ABC theory, the low volume, more complex products made up the bulk of the marginal and unprofitable products" (p. 80).
2. "All of the manufacturing companies explicitly recognized the linkage of ABC to their companies' formal quality initiatives" (p. 275).

3. "A few of the case sites . . . had taken actions based on the insights from their ABC models. Most of the other sites, however, had yet to make a transition from analysis to action" (p. 307).

The authors summarize the intent of their study and its contributions as, "this book represents a state-of-the-art view of ABC principles and implementation experiences of eight companies. The findings will provide an understanding of the design decisions and pitfalls and the organizational dynamics that can directly affect the success of an ABC system" (p. vii). However, as previously indicated, fewer and more detailed case studies would have provided more detailed guidance to practitioners who desire to move toward an ABC implementation on their own. As for academics, this study is definitely worth reading if only for the purpose of stirring up their intellectual juices in preparation for putting together their own research proposal on "Researching the Efficacy of ABC Systems and Their Implementation."

WILLIAM L. FERRARA
Professor of Accountancy
Stetson University

DOUGLAS D. DAVIS and CHARLES A. HOLT, *Experimental Economics* (Princeton: Princeton University Press, 1993, pp. xvi, 572, $39.50).

Experimental economics refers to a research paradigm that uses controlled laboratory conditions to investigate the relative predictive ability of economic theories. This monograph provides an overview of the history and usefulness of economic experimentation, presents a discussion of the controversies regarding the paradigm, and reviews a broad range of experimental economic research. The authors are important contributors to the literature and approach their endeavor with a perspective that is sympathetic to the extant line of experimental inquiry.

The authors begin by providing an example of a double-auction market experiment (where buyers and sellers interact to establish trading prices and final endowments) to discuss the importance of financial incentives and control features in experimental economies. They argue that the primary benefits from experimentation are replicability and control, features necessary for a discipline to build a strong theoretical base.

In Chapter 2 the authors discuss individual decision theory, with a primary focus on tests of two-person game theoretic models. This chapter provides a useful overview of some game theory (particularly the prisoner's dilemma and the battle of the sexes) and a discussion of the usefulness of the backward-induction rationality for calculating equilibria. The major focus of the experimental research has been on the predictive abilility of the Nash equilibria. The authors conclude that the experimental research provides mixed evidence in support for the game theoretic predictions.

Chapter 3 deals with double-auctions, a trading mechanism extensively studied in the laboratory because it has proven so effective and robust in generating outcomes as predicted by economic theories. The research summarized in this chapter investigated how modifying trading rules and environmental features affect market outcomes. The authors highlight several important findings with regard to the double-auction institution, including the result that complete information regarding supply and demand schedules is not necessary to achieve competitive outcomes. Indeed, complete information on the schedules can retard the speed of convergence to the competitive equilibria.

The research reviewed in Chapter 4 employed the posted-offer institution (seller posts a take-it-or-leave-it offer). Again, the discussion focuses on market outcomes under a variety of institutional environmental features. A large portion of this research is framed as addressing industrial organization questions, and the authors suggest caution in over-interpreting the results, but note that this is a rich area for future research.

Research involving bilateral bargaining, single-seller auctions, and two-sided uniform-price auctions is examined in Chapter 5. Of particular interest is a discussion of the winners' curse in various institutions. Chapter 6 deals with research on public goods, externalities, and voting. There is an interesting discussion of incentive compatible mechanisms.

In Chapter 7, the authors discuss asymmetric information. This chapter (along with Chapter 8) may be most interesting to accounting researchers. The discussion covers topics such as the lemons model, reputational effects, signaling, and asset markets. Three conclusions are offered: information asymmetry can materially influence market performance, reputations can overcome some problems of information asymmetry, and the effects of information asymmetry are specific to the trading institution. Accounting researchers have the potential to contribute to this line of research because of their knowledge of information-based institutions and market behavior in the face of information asymmetry.

Chapter 8 provides a fine discussion of the research dealing with individual decision-making in risky situations. The discussion includes commonly addressed issues such as preference reversals, difference between willingness to pay and willingness to accept, representativeness heuristic, curse of knowledge, and framing effects. The authors submit that economists must work to better understand the potential problems of the expected utility hypothesis, the basis for the predictions of individual decision-making. The final chapter of the book provides a well-balanced discussion of experimental methods. The authors argue that the ability to focus on the effect of trading institutions on market performance is a major contribution from experimental economics. The most interesting subsection of this chapter is a discussion of the relation between theoretical, experimental, and natural economic environments.

This book is an excellent reference book for individuals interested in gaining an overall perspective of experimental economics. It allows easy access for individuals considering using the method for research or as a pedagogical tool (the authors provide instructions for a number of different experiments that could be conducted in the classroom). Although the book is not written specifically for an accounting audience, it provides a balanced overview of experimental economics for novices and yet provides a useful review for those more experienced with experimental economic research. The experimental method is used broadly in economic research, and has the potential to be useful to accounting researchers who are addressing problems where access to naturally occurring data is limited. This book could be a useful supplement for introductory Ph.D. courses dealing with design/statistical issues and for theory courses that consider theory-data interactions.

RONALD R. KING
Associate Professor of Accounting

Washington University

PELHAM GORE, *The FASB Conceptual Framework 1973–1985: An Analysis* (Manchester, UK: Manchester University Press, 1992, pp. xi, 189, £69.95).

The Conceptual Framework Project (CF) undertaken by the Financial Accounting Standards Board (FASB) is possibly the most ambitious attempt to establish financial reporting guidelines ever undertaken. The author notes that the project resulted in 60 documents totaling over 4,000 pages. This book is also ambitious. It seeks to (1) provide a historical framework for the CF in order to show why the CF project was undertaken, (2) to determine the role of a CF in general, (3) to discern how the CF works in terms of its impact on practice, and (4) to analyze the international impact of the CF project. Fortunately the book is more successful in attaining its intended purpose than the FASB was in drafting the CF project. As Gore discusses extensively in his book, the CF has been the subject of extensive criticism and litte acclaim. Critics of the CF argue that the project is internally inconsistent, fails to take the economic environment into account, and is of little use to present or future practice. Gore seeks to explain why the CF has not been viewed as successful and to consider the implications of the lack of acceptance of the CF for international standard setting.

The author used a two-pronged approach to evaluating the CF. The book draws from a wide range of documentary evidence including FASB news releases and minutes, academic journal articles, news articles, and articles in practitioner-oriented magazines. The author also conducted extensive interviews with 15 individuals who had been involved in, or closely observed, the development of the CF project.

This book does a credible job of summarizing the history of the accounting standard-setting process in the United States and illustrating the series of events that led to the decision to undertake the CF project. The author argues that the economic environment in which the accounting profession found itself in the late 1960s spawned a crisis that led to a need for the CF project.

In its attempt to evaluate and critique the theoretical issues and problems with respect to the CF the book is less clear. In large part, this is because of the complexity of the process itself and the lack of clear agreement as to what function the CF was ultimately expected to fulfill. The book points out several problems with the CF, the most notable of which seems to be that the FASB did not take into account the political nature of accounting standard setting in the CF document, and yet the CF itself is a political document.

Finally, the author discusses the impact of the CF both in the U.S. and internationally. Within the U.S., the author concludes that the CF has had a limited impact on three areas. First, the CF now serves as a tool used by constituents to argue for or against proposed accounting standards. A difficulty with this, as noted by the author, is that the terminology in the CF is sufficiently ambiguous so as to support widely divergent views. Second, the CF may serve a pedagogical purpose. The author notes that while the CF *could* serve this purpose, it has not been broadly integrated into the U.S. accounting curriculums. Third, the CF has had an effect on accounting standards. The author examines several accounting standards issued after completion of the CF to assess the degree of consistency with the CF (pensions, deferred taxes, and cash flow). The conclusion is that while the CF may have influenced these standards, none are entirely consistent with the CF. In large part, this seems to be due to the failure of the CF to take into account the political environment in which accounting standards are set.

Outside the U.S., the CF has been instrumental in leading other standard setting bodies in other countries as well as the International Accounting Standards Committee to issue their own projects aimed at developing their own conceptual frameworks. By and large, these have been similar in approach and structure to the CF project in the U.S., a point which the author considers unfortunate given the problems with the CF in the U.S.

This book would be a useful addition to a university library as a reference book for researchers interested in a detailed study of the CF project. It would also be valuable reading for standard setters in countries still in the process of developing their own frameworks. By taking into account some of the difficulties with formulation and acceptance of the U.S. conceptual framework, other bodies developing frameworks may be able to learn from, and avoid, some of the difficulties that occurred in the U.S.

FRANCES L. AYRES
Associate Professor

The University of Oklahoma

CHRISTOPHER HUMPHREY, PETER MOIZER, and STUART TURLEY, *The Audit Expectations Gap in the United Kingdom* (London: The Institute of Chartered Accountants in England and Wales, 1992, pp. viii, 99, £10.00, paper).

This report, which was sponsored by the Auditing Research Foundation of The Institute of Chartered Accountants in England and Wales (ICAEW), had four objectives:

> . . . to examine the history of the expectations gap, to collect evidence indicating what kind of gap exists at present, to determine what factors connected with auditing are influential in the formation of expectations, and to suggest ways in which the gap can be closed (p.2).

Although the expression "expectations gap" seems to have been coined in the United States by the Cohen Commission in the 1970s, the authors believe that the history of the audit expectations gap in the United Kingdom dates back to the beginnings of company auditing in the nineteenth century. They suggest that the issues incorporated within the term audit expectations gap (auditors' fraud detection responsibilities, auditor independence, public interest reporting, the meaning of auditors' communications) have not changed much over time. In both the U.K. and the U.S., the response of the auditing profession to this predicament has been to try to educate the public as to the true nature of auditors' duties. The authors believe, however, that the problem is more than just an "ignorance" gap and they conclude their report with three suggestions for bridging the audit expectations gap in the U.K.

American readers of this study will probably find the final chapter of most interest because it is here that the authors discuss the rationale for their three reforms. The first is the establishment of an independent Office for Auditing to oversee the appointment of auditors of large companies and to regulate audit fees. This reform is intended to enhance the independence of auditors from their major clients. The second is to extend the accountability of British auditors to include potential shareholders and existing and potential creditors. At the same time, the authors would like to see the concept of joint and several liability in the U.K. amended to reflect the auditor's degree of responsibility for misstated financial statements. The third suggestion is to make auditors responsible for the detection of material fraud.

The customary disclaimer in this report (p. ii) indicates that the views of the authors are not necessarily shared by the sponsor, the ICAEW. Elsewhere, the ICAEW has expressed a desire for effective reform to be achieved by "a self-regulating profession, working in partnership with government, not by a statutory straitjacket" (*The Economist*, November 21st, 1992, p. 6). Observers outside the profession also have made recommendations to close the expectations gap. For example, a recent editorial in *The Economist* (November 28th, 1992, p. 19) sanctioned the mandatory rotation of auditors and the prohibition against providing nonaudit services such as tax or consulting advice to audit clients. The editorial also called for outside statutory regulation of the auditing profession in the U.K. The editors believe that these changes are necessary if auditors are to regain the confidence of the public after several recent, highly-publicized audit failures in Britain (BCCI) and in America (thrifts).

Though the report under review offers several solutions to a complex problem that has been said to exist in both the U.K. and the U.S. for a very long time, this reviewer believes that it may be premature to attempt to implement any of them. What the profession needs to do first is to obtain hard evidence on the magnitude of the audit expectations gap and the costs and benefits of the various remedies that have been proposed in this study and elsewhere. What is the proportion of audits that are substandard? Is the ratio increasing or decreasing? Is the ratio approximately the same around the world, or is it higher in some areas than others? What are the costs to investors and to societies of audit failure? How much do auditors earn from company audits and from nonaudit work? It is unfortunate that the study reviewed here did not attempt to provide answers to these and related questions, for if it had, it could have made a significant contribution to the debate on how to more effectively regulate the auditing profession in the U.K., the U.S., and elsewhere in the world.

STEPHEN TOMCZYK
Associate Professor of Accounting
Suffolk University

YUJI IJIRI and ISAO NAKANO, editors, *Business Behavior and Information* (Pittsburgh: Carnegie Mellon University Press, 1992, pp. xvi, 200, $15.00, paper).

This book consists of seven papers presented at the Kobe forum (held in Kobe, Japan) with the same title as the book, along with a brief introduction by the editors. The conference organizers seem to have adopted a very broad definition of business behavior. The advantage to this approach is that it helps the reader break out of the traditional paradigms and consider different viewpoints of the impact of information on business behavior. For example, one paper argues that we may best understand firms as a collection of "information fields" or *ba*'s, similar to the agency view of the firm as a nexus of contracts. The downside to a broad definition of business behavior is that each of the seven papers adopts a strikingly different perspective toward the problem. As a result, the potential for synergies created by focusing on a single, narrowly defined topic (or methodology) is lost. The editors, however, have done an excellent job in providing a coherent framework within which the papers can be related.

The papers address business behavior at the macro, micro, and intermediate levels. For example, the paper by Professor Shyam Sunder adopts a macro perspective and examines insider trading as a tradeoff between alleviating information asymmetrics in the market for a firm's securities and protecting the shareholders' property rights. Examining how information affects relationships within the firm, Professor Yoshihara's paper advocates the increased use of English by Japanese multinationals as a way to truly internationalize their operations. Similarly, Professor Ijiri's paper highlights the interesting point that mere knowledge that verifiable information will be available is enough to affect behavior and, hence, detailed performance-related information may be explicitly used only in "off-equilibrium" paths. At the micro, individual level, Professors McGuire and Malone present the results of a behavioral study to argue that individual decision makers discount both base rates and case cues.

Two other points are worthy of note. First, while the unit of analysis, methodology, and rigor vary across papers, most papers have the implicit notion that information is best utilized (in the sense of maximizing an implicit social welfare function or obtaining a Pareto improvement) in a cooperative, "team-like" setting. Second, although each paper had a listed commentator, their comments are not included in the book. I would have liked access to the discussant's comments after each paper because, quite often, such comments provide a differing viewpoint and, thus, enhance the reader's learning.

In sum, this book brings together a variety of perspectives which serve to highlight the various uses of information, within and across business entities. Hence, it is likely to make a useful addition to a business library.

RAMJI BALAKRISHNAN
Assistant Professor of Accounting
The University of Iowa

MARILYN KLEINBERG NEIMARK, *The Hidden Dimensions of Annual Reports: Sixty Years of Social Conflict at General Motors* (New York: Markus Wiener Publishing, Inc., 1992, pp. v, 207, $34.95).

With *The Hidden Dimensions of Annual Reports*, Marilyn Kleinberg Neimark helps to fill a vacuum in the accounting literature. The book provides a useful and understandable exposition of an alternative theory for explaining organizational change, an expanded perspective on what constitutes "accounting," and an informative critique of accounting's role in some of our current economic problems. Neimark's book would add depth to accounting courses concerned with accounting theory or the responsibilities of accountants.

Two chapters of the book will be particularly useful for graduate accounting theory courses. Chapter 2 provides an easy-to-understand and trenchant critique of transaction cost theory. Contrary to advocates of transaction cost theory who maintain that the theory is based on the "neutral" concept of efficiency, Neimark exposes the ideological underpinnings of this approach. In Chapter 3, she describes the dialectical approach to organizational change, which challenges some of the fundamental assumptions underlying much accounting research. What makes this chapter particularly valuable is that Neimark avoids the jargon frequently seen in this type of work; and when she does use unfamiliar terms (such as "dialectical" and "ideology"), she provides definitions.

The main section of the book uses the dialectical theory to analyze the non-financial aspects of General Motors' annual reports (Chaps. 4, 5, and 6). Neimark illustrates the manner in which the verbal portions of the reports are used to bolster the goals of GM management. Throughout the history of the company, Neimark argues, the reports were used to assert the superiority of management over the unions. In addition, she demonstrates how GM used the annual reports to support social consumption, reinforce "family values," magnify the importance of technology, and justify moving production facilities to other countries. While these findings may not be surprising, Neimark's exposition and theoretical framework provide a new way to think about the role of these reports; she emphasizes the fact that accounting consists not merely of financial statements, but also includes the photos and text of annual reports—what she calls their "Hidden Dimensions."

The other chapters in the book also illustrate dialectical theory. There is a fascinating and highly readable history of the automobile industry from the labor perspective (Chaps. 1 and 4). These chapters would contribute to any business history or labor relations class. The introduction and conclusion provide a perspective on current conditions in the industry and the United States economy as a whole. They address Japanese competition, quality concerns, the unequal distribution of wealth, and the budget deficit, among other topics. In the conclusion, Neimark discusses the role accountants have played in these areas and suggests some seldom-heard reforms, including breaking the financial links between corporations and their auditors.

I strongly recommend that this book be considered by those who teach accounting theory, business history, labor relations, accounting ethics, or contemporary issues in accounting. Even if—no, *especially* if—you doubt you would agree with her perspective, it is an important book. The book would provide students with an alternative perspective, thus helping them develop critical analytical skills. For similar reasons, the book could also be useful to accounting academics who want to give careful consideration to their theoretical orientations.

THERESA D. HAMMOND
Assistant Professor of Accounting
Boston College

THOMAS F. POGUE, editor, *State Taxation of Business: Issues and Policy Options* (Westport, CT: Praeger Publishers, 1992, pp. xviii, 340, $55.00).

Until recently, state tax policy was not a major concern of accountants or accounting scholars. Academic interest in the area generally focused on Constitutional issues, such as the impact of state taxation on interstate commerce, or First Amendment "rights" to statewide (versus local) funding of public schools. State tax policy thus fell within the purview of lawyers and law schools, rather than accountants and business schools. Even tax practitioners gave the areas short shrift—logically so, because state tax burdens (net of Federal tax deduction benefits) tended to be small relative to the Federal burden; therefore, the high cost of planning for a multitude of state taxes could be expected to outweigh its benefits.

Federal tax reform in the early 1980s, combined with state fiscal crises of the late 1980s, changed this. The halving of top Federal income tax rates, coupled with cutbacks (both in itemized deductions and the alternate minimum tax) in Federal deductibility along with steep increases in state tax rates dramatically increased the ratio of state and local taxes to total tax burden. The overall increase in the relative burden of state and local taxes awakened interest among business scholars in the wide ranging economic impacts of state taxes. Indeed, some downright brilliant work has been published over the past few years on the impact of state taxes on siting decisions (that is, decisions by both foreigners and Americans about locations for expanding into new facilities or contracting into one of several old ones).

State Taxation of Business does a yeoman's job of identifying the most important of the economic impacts of state and local taxation in the United States. This is not surprising, as the book developed out of the National Tax Association's 1991 conference on the subject. A wide range of views on policy issues, and policy options for dealing with such issues, is presented in the book's 20 chapters (each of which grew out of a paper presented at the conference). Although none of the papers has been edited so as to unify the book, a theme emerges from the disparate positions proffered by the nearly 30 authors.

If the papers agglomerated into *State Taxation of Business* agree on anything, it is that state and local taxation in the United States is not so much a system based on the strategic application of over-arching policies as it is a disjointed product of tactical, step-wise reactions to per ceived taxpayer activities. There is no agreement, and seems to have been little governmental thought, as to how states should best tax business operations.

What the papers make clear, however, is that whatever tax policies might have been optimal in the past likely will not be optimal in the near future. Not so much because of clever tax planning or pernicious tax fraud, but instead due to fundamental changes in the business world. Changes wrought by the information revolution in the ways (and places) businesses basically do business as well as by changes in the competitive position of states in an increasingly globalized economy. States need to reformulate tax policies, not just to reform tax systems. *State Taxation* *of Business* comprehensively presents the issues emerging from the great laboratory of the many independent jurisdictions which characterize state and local taxation in the United States. As such, the book is worthy of attention by all those engaged in researching, formulating, or implementing tax policy.

JOHN E. KARAYAN
Adjunct Professor of Accounting
University of Southern California

BOOKS RECEIVED FEBRUARY 5, 1993 TO MAY 1, 1993

RAJ AGGARWAL, editor, *Capital Budgeting Under Uncertainty* (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1993, pp. x, 292, $53.33).

KWABENA ANYANE-NTOW, editor, *International Handbook of Accounting Education and Certification* (Oxford, UK: Pergamon Press Ltd., 1992, pp. xxxii, 556, £65.00).

R. C. BILLINGS, et al., *Environmental Auditing and the Role of the Accounting Profession* (Toronto: The Canadian Institute of Chartered Accountants, 1992, pp. xiv, 156, C$34.50, paper).

JOHN M. BOERSEMA and SUSAN J. VAN WEELDEN, *Financial Reporting for Segments* (Toronto: The Canadian Institute of Chartered Accountants, 1992, pp. xiv, 237, C$39.50, paper).

STEPHEN A. BUTLER, editor, *Earnings Quality* (Norman, OK: The Center for Economic and Management Research, 1993, pp. iii, 183, $15.00, paper).

CAROLE B. CHEATHAM and LEO R. CHEATHAM, *Updating Standard Cost Systems* (Westport, CT: Quorum Books, 1993, pp. xiv, 223, $59.95).

GARY COKINS, ALAN STRATTON, and JACK HELBLING, *An ABC Manager's Primer* (Montvale, NJ: Institute of Management Accountants, 1993, pp. xii, 65, $15.00, paper).

COOPERS & LYBRAND, *1993 International Tax Summaries: A Guide for Planning and Decisions* (New York: John Wiley & Sons, Inc., 1993, pp. x, 1,302, $105.00).

MARC J. EPSTEIN and ALBERT D. SPALDING, JR., *The Accountant's Guide to Legal Liability and Ethics* (Homewood, IL: Richard D. Irwin, Inc., 1995, pp. xviii, 296, $17.50, paper).

DOUGLAS GARBUTT, *Making Budgets Work: The Control and Use of the Budgetary Control Process* (London: The Chartered Institute of Management Accountants, 1992, pp. xv, 182, £17.50, paper).

JAMES C. HUMES, *The Sir Winston Method: The Five Secrets of Speaking the Language of Leadership* (New York: William Morrow and Company, Inc., 1991, pp. 189, $7.00, paper).

YASUHIRO MONDEN, *Toyota Management System: Linking the Seven Key Functional Areas* (Cambridge, MA: Productivity Press, Inc., 1993, pp. xxiii, 222, $45.00).

KENNETH N. MYERS, *Total Contingency Planning for Disasters* (New York: John Wiley & Sons, Inc., 1993, pp. xvi, 270, $55.00).

PAUL PACTER, *Reporting Disaggregated Information* (Norwalk, CT: Financial Accounting Standards Board, 1993, pp. xxii, 423, $25.00 non-members, $20.00 members/academics, paper).

THOMAS F. POGUE, editor, *State Taxation of Business: Issues and Policy Options* (Westport, CT: Praeger Publishers, 1992, pp. xviii, 340, $55.00).

AHMED RIAHI-BELKAOUI, *Quality and Control* (Westport, CT: Quorum Books, 1993, pp. xi, 219, $59.95).

MORDEN S. SHAPIRO, *Mergers of Professional Practices: Managing the Process* (Toronto: The Canadian Institute of Chartered Accountants, 1992, pp. xvii, 164, C$49.50, paper).

CRISPIN SIMON, *The Role of the Accountant in Strategic Planning* (Brookfield, VT: Gower Publishing Company Limited, 1992, pp. x, 116, $49.95).

J. DAVID SPICELAND and SURENDRA AGRAWAL, editors, *International Guide to Accounting Journals, Second Edition* (Princeton: Markus Wiener Publishing, Inc., 1993, pp. vi, 314, $49.95).

HERBERT R. STEINBACHER and NORMA L. STEINBACHER, *TPM for America: What It Is and Why You Need It* (Cambridge, MA: Productivity Press, 1993, pp. xxii, 147, $20.00).

ROGER Y. W. TANG, *Transfer Prices in the 1990s: Tax and Management Perspectives* (Westport, CT: Quorum Books, 1993, pp. xviii, 143, $45.00).

ROBERT J. THIERAUF, *Creative Computer Software for Strategic Thinking and Decision Making: A Guide for Senior Management and MIS Professionals* (Westport, CT: Quorum Books, 1993, pp. xix, 321, $49.95).

GARY TIDWELL, *Anatomy of a Fraud: Inside the Finances of the PTL Ministries* (New York: John Wiley & Sons, Inc., 1993, pp. viii, 357, $24.95).

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal.... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow (with a slight modification) the B-format of the *Chicago Manual of Style* (13th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's International Dictionary*.

FORMAT

1. All manuscripts should be typed on one side of 8½ × 11" good quality paper and be double spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To assure anonymous review, authors should not identify themselves directly or indirectly in their papers. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title, and affiliation, any acknowledgements, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: *All* pages, including tables, appendices, and references, should be serially numbered. The first section of the paper should be untitled and unnumbered. Major sections may be numbered in roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights, and measures. For example: *three days*; *3 kilometers*; *30 years*. All other numbers are expressed numerically. Generally when using approximate terms spell out the number, for example, *approximately thirty years*.

Percentages and Decimal Fractions: In nontechnical copy use the word *percent* in the text; in technical copy the symbol % is used. (See the *Chicago Manual* for discussion of these usages.)

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: a *well-presented analysis*; *re-form*. See *Webster's* for correct usage.

Key Words: The synopsis is to be followed by four *key words* that will assist in indexing the paper.

SYNOPSIS

A synopsis of about 400 words should be presented on a separate page immediately preceding the text. The synopsis also serves as the "introduction." The synopsis should be nonmathematical and include a readable summary of the research question, method, and the significance of the findings and contribution. The title, but not the author's name or other identification designations, should appear on the synopsis page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an arabic number and a complete title indicating the exact contents of the table or figure.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix *a*, *b*, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says. . . ."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by the *Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Dates of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) should be listed in chronological order of publication. Two or more works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 8.67.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2d ed. Homewood, IL: Dow Jones-Irwin.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, UK: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–36.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, University of Texas, Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be double spaced and numbered consecutively throughout the manuscript with superscript Arabic numerals. Footnotes are placed at the end of the text.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or other publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan, or the like) should be submitted.
3. Four copies should be submitted together with a check in U.S. funds for $50.00 for members or $100.00 for nonmembers of the AAA made payable to the American Accounting Association. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the Editor for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting.

"An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional, research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

THE ACCOUNTING REVIEW

Index to Volume 68 1993

Published Quarterly by the
AMERICAN ACCOUNTING ASSOCIATION

AUTHOR INDEX TO VOLUME 68

Author	*Title*	*Page*
ALCIATORE, MIMI L.	New Evidence on SFAS No. 69 and the Components of Change in Reserve Value	639
ALLES, MICHAEL (and LUNDHOLM)	On the Optimality of Public Signals in the Presence of Private Information	93
AMIR, ELI	The Market Valuation of Accounting Information: The Case of Postretirement Benefits other than Pensions	703
BABAD, YAIR (and BALACHANDRAN)	Cost Driver Optimization in Activity-Based Costing	563
BAGINSKI, STEPHEN P. (and CONRAD and HASSELL)	The Effects of Management Forecast Precision on Equity Pricing and on the Assessment of Earnings Uncertainty	913
BALACHANDRAN, BALA V. (and BABAD)	Cost Driver Optimization in Activity-Based Costing	563
BALL, RAY (and KOTHARI and WATTS)	Economic Determinants of the Relation Between Earnings Changes and Stock Returns	622
BANKER, RAJIV D. (and JOHNSTON)	An Empirical Study of Cost Drivers in the U.S. Airline Industry	576
BANKER, RAJIV D. (and DAS and DATAR)	Complementarity of Prior Accounting Information: The Case of Stock Dividend Announcements	28
BARTOV, ELI	The Timing of Asset Sales and Earnings Manipulation	840
BENEISH, MESSOD D. (and PRESS)	Costs of Technical Violation of Accounting-Based Debt Covenants	233
BROWN, LAWRENCE D. (and KIM)	The Association Between Nonearnings Disclosures by Small Firms and Positive Abnormal Returns	668
BUCKLESS, FRANK A. (and PEACE)	The Influence of the Source of Professional Standards on Juror Decision Making	164
BUSHMAN, ROBERT M. (and INDJEJIKIAN)	Stewardship Value of "Distorted" Accounting Disclosures	765
CALLEN, JEFFREY L. (and FALK)	Agency and Efficiency in Nonprofit Organizations: The Case of "Specific Health Focus" Charities	48
CHEN, KEVIN C. W. (and WEI)	Creditors' Decisions to Waive Violations of Accounting-Based Covenants	218
CHRIST, MARY YORK	Evidence on the Nature of Audit Planning Problem Representations: An Examination of Auditor Free Recalls	304
CHUNG, KWANG-HYUN (and GHICAS and PASTENA)	Lenders' Use of Accounting Information in the Oil and Gas Industry	885
CONRAD, EDWARD J. (and BAGINSKI and HASSELL)	The Effects of Management Forecast Precision on Equity Pricing and on the Assessment of Earnings Uncertainty	913
CROWE, JOHN (and TUNG)	Expectation Formation and Financial Ratio Adjustment Processes: A Comment and An Extension	942
DARROUGH, MASAKO N.	Disclosure Policy and Competition: Cournot vs. Bertrand	534
DAS, SOMNATH (and BANKER and DATAR)	Complementarity of Prior Accounting Information: The Case of Stock Dividend Announcements	28
DATAR, SRIKANT M. (and KEKRE and MUKHOPADHYAY and SRINIVASAN)	Simultaneous Estimation of Cost Drivers	602

Author	*Title*	*Page*
DATAR, SRIKANT M. (and BANKER and DAS)	Complementarity of Prior Accounting Information: The Case of Stock Dividend Announcements	28
DAVIS, HARRY ZVI (and PELES)	Measuring Equilibrating Forces in Financial Ratios	725
DAVIS, JON S. (and SWENSON)	Experimental Evidence on Tax Incentives and the Demand for Capital Investments	482
DAVIS, LARRY R. (and RICCHIUTE and TROMPETER)	Audit Effort, Audit Fees, and the Provision of Nonaudit Services to Audit Clients	135
DONTOH, ALEX (and RONEN)	Information Content of Accounting Announcements	857
DOPUCH, NICHOLAS	A Perspective on Cost Drivers	615
DUNK, ALAN S.	The Effect of Budget Emphasis and Information Asymmetry on the Relation Between Budgetary Participation and Slack	400
EL-GAZZAR, SAMIR M.	Stock Market Effects of the Closeness to Debt Covenant Restrictions Resulting from Capitalization of Leases	258
ENIS, CHARLES R.	Earnings-Based Compensation Plan, Performance, and Capital Expenditure Policy in the Motor Carrier Industry	928
FALK, HAIM (and CALLEN)	Agency and Efficiency in Nonprofit Organizations: The Case of "Specific Health Focus" Charities	48
FISHER, JOSEPH (and YOUNG and LINDQUIST)	The Effects of Intergroup Competition and Intragroup Cooperation on Slack and Output in a Manufacturing Setting	466
FLORY, STEVEN M. (and PHILLIPS, JR. and REIDENBACH and ROBIN)	A Reply to "A Comment on 'A Multidimensional Analysis of Selected Ethical Issues in Accounting' "	417
FOSTER III, TAYLOR W. (and KOOGLER and VICKREY)	Valuation of Executive Stock Options and the FASB Proposal: An Extension	184
GHICAS, DIMITRIOS (and CHUNG and PASTENA)	Lenders' Use of Accounting Information in the Oil and Gas Industry	885
HANSEN, STEPHEN C.	Strategic Sampling, Physical Units Sampling, and Dollar Units Sampling	323
HASSELL, JOHN M. (and BAGINSKI and CONRAD)	The Effects of Management Forecast Precision on Equity Pricing and on the Assessment of Earnings Uncertainty	913
HAYES, ROBERT D. (and MILLAR)	A Rejoinder to "Measuring Production Efficiency in a Not-for-Profit Setting: An Extension"	89
HO, LI-CHIN JENNIFER	Option Trading and the Relation Between Price and Earnings: A Cross-Sectional Analysis	368
INDJEJIKIAN, RAFFI J. (and BUSHMAN)	Stewardship Value of "Distorted" Accounting Disclosures	765
JENNERGREN, L. PETER (and NASLUND)	A Comment on "Valuation of Executive Stock Options and the FASB Proposal"	179
JOHNSTON, HOLLY H. (and BANKER)	An Empirical Study of Cost Drivers in the U.S. Airline Industry	576
JONES, SCOTT K. (and PONEMON)	A Comment on "A Multidimensional Analysis of Selected Ethical Issues in Accounting"	411
KEKRE, SUNDER (and DATAR and MUKHOPADHYAY and SRINIVASAN)	Simultaneous Estimation of Cost Drivers	602
KIDA, THOMAS (and MALETTA)	The Effect of Risk Factors on Auditors' Configural Information Processing	681

Author	*Title*	*Page*
KIM, KWON-JUNG (and BROWN)	The Association Between Nonearnings Disclosures by Small Firms and Positive Abnormal Returns	668
KOOGLER, PAUL R. (and FOSTER III and VICKREY)	Valuation of Executive Stock Options and the FASB Proposal: An Extension	184
KOTHARI, S. P. (and BALL and WATTS)	Economic Determinants of the Relation Between Earnings Changes and Stock Returns	622
LANGER, RUSSELL (and LEV)	The FASB's Policy of Extended Adoption for New Standards: An Examination of FAS No. 87	515
LEE, CHENG F. (and WU)	Expectation Formation and Financial Ratio Adjustment Processes: A Reply	953
LEV, BARUCH (and LANGER)	The FASB's Policy of Extended Adoption for New Standards: An Examination of FAS No. 87	515
LI, SHU-HSING (and MENSAH)	Measuring Production Efficiency in a Not-for-Proft Setting: An Extension	66
LINDQUIST, TIM M. (and YOUNG and FISHER)	The Effects of Intergroup Competition and Intragroup Cooperation on Slack and Output in a Manufacturing Setting	466
LIPE, MARLYS GASCHO	Analyzing the Variance Investigation Decision: The Effects of Outcomes, Mental Accounting, and Framing	748
LOREK, KENNETH S. (and SCHAEFER and WILLINGER)	Time-Series Properties and Predictive Ability of Funds Flow Variables	151
LUNDHOLM, RUSSELL (and ALLES)	On the Optimality of Public Signals in the Presence of Private Information	93
MALETTA, MARIO J. (and KIDA)	The Effect of Risk Factors on Auditors' Configural Information Processing	681
MCMILLAN, JEFFREY J. (and WHITE)	Auditors' Belief Revisions and Evidence Search: The Effect of Hypothesis Frame, Confirmation Bias, and Professional Skepticism	443
MENSAH, YAW M. (and LI)	Measuring Production Efficiency in a Not-for-Profit Setting: An Extension	66
MILLAR, JAMES A. (and HAYES)	A Rejoinder to "Measuring Production Efficiency in a Not-for-Profit Setting: An Extension"	89
MITTELSTAEDT, H. FRED (and REGIER)	The Market Response to Pension Plan Terminations	1
MOHRMAN, MARY BETH	Debt Contracts and FAS No. 19: A Test of the Debt Covenant Hypothesis	289
MORTON, SANFORD	Strategic Auditing for Fraud	825
MUKHOPADHYAY, TRIDAS (and DATAR and KEKRE and SRINIVASAN)	Simultaneous Estimation of Cost Drivers	602
MUTCHLER, JANE F. (and PANY)	A Comment on "The Influence of the Source of Professional Standards on Juror Decision Making"	176
NASLUND, BERTIL (and JENNERGREN)	A Comment on "Valuation of Executive Stock Options and the FASB Proposal"	179
NELSON, MARK W.	The Effects of Error Frequency and Accounting Knowledge on Error Diagnosis in Analytical Review	804
PANY, KURT (and MUTCHLER)	A Comment on "The Influence of the Soruce of Professional Standards on Juror Decision Making"	176
PARKASH, MOHINDER (and VENABLE)	Auditee Incentives for Auditor Independence: The Case of Nonaudit Services	113
PASTENA, VICTOR (and CHUNG and GHICAS)	Lenders' Use of Accounting Information in the Oil and Gas Industry	885
PEACE, ROBERT L. (and BUCKLESS)	The Influence of the Source of Professional Standards on Juror Decision Making	164

Author	*Title*	*Page*
PELES, YORAM C. (and DAVIS)	Measuring Equilibrating Forces in Financial Ratios	725
PHILLIPS, JR., THOMAS J. (and FLORY and REIDENBACH and ROBIN)	A Reply to "A Comment on 'A Multidimensional Analysis of Selected Ethical Issues in Accounting' "	417
PONEMON, LAWRENCE A. (and JONES)	A Comment on "A Multidimensional Analysis of Selected Ethical Issues in Accounting"	411
POWNALL, GRACE (and WASLEY and WAYMIRE)	The Stock Price Effects of Alternative Types of Management Earnings Forecasts	896
PRESS, ERIC (and BENEISH)	Costs of Technical Violation of Accounting-Based Debt Covenants	233
REGIER, PHILIP R. (and MITTELSTAEDT)	The Market Response to Pension Plan Terminations	1
REIDENBACH, R. ERIC (and FLORY and PHILLIPS, JR. and ROBIN)	A Reply to "A Comment on 'A Multidimensional Analysis of Selected Ethical Issues in Accounting' "	417
RICCHIUTE, DAVID N. (and DAVIS and TROMPETER)	Audit Effort, Audit Fees, and the Provision of Nonaudit Services to Audit Clients	135
ROBIN, DONALD P. (and FLORY and PHILLIPS, JR. and REIDENBACH)	A Reply to "A Comment on 'A Multidimensional Analysis of Selected Ethical Issues in Accounting' "	417
RONEN, JOSHUA (and DONTOH)	Information Content of Accounting Announcements	857
SANSING, RICHARD C.	Information Acquisition in a Tax Compliance Game	874
SCHAEFER, THOMAS F. (and LOREK and WILLINGER)	Time-Series Properties and Predictive Ability of Funds Flow Variables	151
SHAW, WAYNE H. (and WIER)	Organizational Form Choice and the Valuation of Oil and Gas Producers	657
SMITH, JR., CLIFFORD W.	A Perspective on Accounting-Based Debt Covenant Violations	289
SRINIVASAN, KANNAN (and DATAR and KEKRE and MUKHOPADHYAY)	Simultaneous Estimation of Cost Drivers	602
STRONG, NORMAN (and WALKER)	The Explanatory Power of Earnings for Stock Returns	385
SWENSON, CHARLES W. (and DAVIS)	Experimental Evidence on Tax Incentives and the Demand for Capital Investments	482
TEOH, SIEW HONG (and WONG)	Perceived Auditor Quality and the Earnings Response Coefficient	346
TROMPETER, GREG (and DAVIS and RICCHIUTE)	Audit Effort, Audit Fees, and the Provision of Nonaudit Services to Audit Clients	135
TUNG, SAMUEL (and CROWE)	Expectation Formation and Financial Ratio Adjustment Processes: A Comment and An Extension	942
VENABLE, CAROL F. (and PARKASH)	Auditee Incentives for Auditor Independence: The Case of Nonaudit Services	113
VERRECCHIA, ROBERT E.	How Do We Assess a Model of Price and Volume?	870

Author	*Title*	*Page*
WALKER, MARTIN (and STRONG)	The Explanatory Power of Earnings for Stock Returns	385
WALLER, WILLIAM S.	Auditors' Assessments of Inherent and Control Risk in Field Settings	783
WASLEY, CHARLES (and POWNALL and WAYMIRE)	The Stock Price Effects of Alternative Types of Management Earnings Forecasts	896
WATTS, ROSS L. (and BALL and KOTHARI)	Economic Determinants of the Relation Between Earnings Changes and Stock Returns	622
WAYMIRE, GREGORY (and POWNALL and WASLEY)	The Stock Price Effects of Alternative Types of Management Earnings Forecasts	896
WEI, K. C. JOHN (and CHEN)	Creditors' Decisions to Waive Violations of Accounting-Based Debt Covenants	218
WHITE, RICHARD A. (and MCMILLAN)	Auditors' Belief Revisions and Evidence Search: The Effect of Hypothesis Frame, Confirmation Bias, and Professional Skepticism	443
WIER, HEATHER A. (and SHAW)	Organizational Form Choice and the Valuation of Oil and Gas Producers	657
WILLINGER, G. LEE (and LOREK and SCHAEFER)	Time-Series Properties and Predictive Ability of Funds Flow Variables	151
WONG, T. J. (and TEOH)	Perceived Auditor Quality and the Earnings Response Coefficient	346
WU, CHUNCHI (and LEE)	Expectation Formation and Financial Ratio Adjustment Processes: A Reply	953
YOUNG, S. MARK (and FISHER and LINDQUIST)	The Effects of Intergroup Competition and Intragroup Cooperation on Slack and Ouput in a Manufacturing Setting	466

BOOK REVIEWS

Author	*Title*	*Reviewer*	*Pag*
Aggarwal, Raj	Capital Budgeting Under Uncertainty	SUSAN S. HAMLEN	95
Allen, David Grayson, and Kathleen McDermott	Accounting Success: A History of Price Waterhouse in America 1890–1990	R. GLEN BERRYMAN	95
Anyane-Ntow, Kwabena	International Handbook of Accounting Education and Certification	FELIX POMERANZ	95
Barnes, D. M.	Value-for-Money Audit Evidence	PAUL A. JANELL	42
Bavishi, Vinod B.	An Analysis of the Financial Performance of the World's Leading 7,500 Companies	HELEN GERNON	42
Boys, Peter, and John Freear	Accounting History 1976–1986: An Anthology	HANS. V. JOHNSON	42
Bruns, Jr., William	Performance Measurement, Evaluation, and Incentives	MARVIN L. BOUILLON	42
Chambers, R. J.	Foundations of Accounting	DALE A. BUCKMASTER	20
Cooke, T. E., and M. Kikuya	Financial Reporting in Japan: Regulation, Practice and Environment	HELEN GERNON	42
Cooper, Robin, Robert S. Kaplan, Lawrence S. Maisel, Eileen Morrissey, and Ronald N. Oehn	Implementing Activity-Based Cost Management: Moving from Analysis to Action	WILLIAM L. FERRARA	95
Davis, Douglas D., and Charles A. Holt	Experimental Economics	RONALD R. KING	96

Author	*Title*	*Reviewer*	*Page*
Davia, Howard R., Patrick C. Coggins, John C. Wideman, and Joseph T. Kastantin	Managment Accountant's Guide to Fraud Discovery and Control	JAMES W. BRACKNER	206
Dean, Graeme W., Marc P. Yoye, and Paul J. Blayney	Strategic Management Accounting Survey: Overhead Cost Allocation & Performance Evaluation Practices of Australian Manufacturers	SHU S. LIAO	426
Donaldson, John	Business Ethics: A European Casebook	MARY GREENAWALT	207
Enthoven, Adolf J. H., Jaroslav V. Sokolov and Alexander M. Petrachkov	Doing Business in Russia and the Other Former Soviet Republics: Accounting and Joint Venture Issues	TIMOTHY S. DOUPNIK	427
Fay, Jack R.	Accounting Certification, Educational & Reciprocity Requirements: An International Guide	PAUL E. BAYES	692
Fitzgerald, Lin, Robert Johnston, Stan Brignall, Rhian Silvestro, and Christopher Voss	Performance Measurement in Service Businesses	RONALD J. LEWIS	207
Gore, Pelham	The FASB Conceptual Framework 1973–1985: An Analysis	FRANCES L. AYRES	961
Hoogendoorn, Martin N., and Lilliane B. Bijl	Financial Reporting in the Nineties: Regulation and Innovation at a Crossroads	ADOLF J. H. ENTHOVEN	428
Humphrey, Christopher, Peter Moizer, and Stuart Turley	The Audit Expectations Gap in the United Kingdom	STEPHEN TOMCZYK	962
Ijiri, Yuji, and Isao Nakano	Business Behavior and Information	RAMJI BALAKRISHNAN	963
Lauer, Thomas W., Eileen Peacock, and Arthur C. Graesser	Questions and Information Systems	MARTHA M. EINING	208
Lehman, Cheryl R.	Accounting's Changing Role in Social Conflict	CRISTI K. LINDBLOM	693
Lorensen, Leonard	Accounting for Liabilities	EILEEN PEACOCK	430
McKinnon, Sharon M., and William J. Bruns, Jr.	The Information MOSAIC	FRANK R. PROBST	431
Monden, Yasuhiro	Cost Management in the New Manufacturing Age: Innovation in the Japanese Automotive Industry	HELEN M. TRAUGH	432
Neimark, Marilyn Kleinberg	The Hidden Dimensions of Annual Reports: Sixty Years of Social Conflict at General Motors	THERESA D. HAMMOND	964
Pogue, Thomas F.	State Taxation of Business: Issues and Policy Options	JOHN E. KARAYAN	965
Rahman, Asheq Razaur	The Australian Accounting Standards Review Board: The Establishment of its Participative Review Process	W. T. WREGE	433

Author	*Title*	*Reviewer*	*Page*
Rosenthal, Robert	Meta-Analytic Procedures for Social Research	SOON-YONG KWON	434
Rudestan, Kjell Erik, and Rae R. Newton	Surviving Your Dissertation: A Comprehensive Guide to Content and Process	GERALD L. SALAMON	694
Schweitzer, Marcell, Ernst Trossman, and Gerald H. Lawson	Break-Even Analysis: Basic Model, Variants, Extensions	MAURICE L. HIRSCH, JR.	209
Spencer, M. Maxwell	The How & Why of a Tax on Wealth	DALE BANDY	434
Steele, Anthony	Audit Risk and Audit Evidence: The Bayesian Approach to Statistical Auditing	RICHARD A. GRIMLUND	210
Wallace, Wanda A., and S. Duane Smith	Accounting for Options: No Longer an Option	JAMES F. WAEGELEIN	695
Weetman, Pauline	SSAP 15: Accounting for Deferred Taxation	ROBERT M. KOZUB	695
Weetman, Pauline, and Sidney Gray	SSAP 19: Accounting for Investment Properties	JOHN T. RIGSBY	696
Wei, Tang Yun, Lynne Chow, and Barry J. Cooper	Accounting and Finance in China: A Review of Current Practice	ROBERT A. LYON	436
Zeff, Stephen A., Franz Van Der Wel, and Kees Camfferman	Company Financial Reporting: A Historical and Comparative Study of the Dutch Regulatory Process	NORLIN G. RUESCHHOFF	697

BOO!

OK. WE REALIZE THAT YOU PROBABLY WEREN'T TOO FRIGHTENED BY OUR HEADLINE. BUT AT IRWIN, OUR MAIN GOAL AROUND OCTOBER 31ST IS NOT TO TRICK YOU, BUT TO TREAT YOU TO THE MOST UP-TO-DATE ACCOUNTING TEXTS AND SUPPLEMENTS ON THE MARKET FOR YOUR COURSES. BECAUSE FACING A CLASSROOM OF CONFUSED STUDENTS IS NOT UNLIKE A DENTIST FACING A BAND OF RABID TRICK-OR-TREATERS WITH A TASTE FOR CHOCOLATE – *AND ALL HE'S GIVING OUT IS DENTAL FLOSS.*

Anthony and Young
MANAGEMENT CONTROL IN NON-PROFIT ORGANIZATIONS
Fifth Edition

Bernstein and Maksy
CASES IN FINANCIAL STATEMENT REPORTING AND ANALYSIS
Second Edition

Boatsman, Griffin, Vickrey, and Williams
ADVANCED ACCOUNTING
Seventh Edition

Brownlee, Ferris, and Haskins
CORPORATE FINANCIAL REPORTING
Text and Cases
Second Edition

Danos and Imhoff
INTRODUCTION TO FINANCIAL ACCOUNTING
Second Edition

Engstrom and Hay
ESSENTIALS OF NON-PROFIT ACCOUNTING FOR PUBLIC ADMINISTRATORS

Garrison and Noreen
MANAGERIAL ACCOUNTING
Seventh Edition

Hermanson, Walker, Plunkett, and Turner
COMPUTERIZED ACCOUNTING WITH PEACHTREE COMPLETE ACCOUNTING® VERSION 6.0

Hopson and Meyer
INCOME TAX FUNDAMENTALS FOR 1993 TAX RETURNS
1994 Edition

Hoyle
ADVANCED ACCOUNTING
Fourth Edition

Larson, Spoede, and Miller
FUNDAMENTALS OF FINANCIAL AND MANAGERIAL ACCOUNTING

Maher and Deakin
COST ACCOUNTING
Fourth Edition

Miller, Redding, and Bahnson
THE FASB
The People, the Process, and the Politics

Mueller, Gernon, and Meek
ACCOUNTING: AN INTERNATIONAL PERSPECTIVE
Third Edition

Pany and Whittington
AUDITING

Paul
APPLE BLOSSOM COLOGNE
Fourth Edition

Pratt and Kulsrud
THE PRATT TAXATION SERIES
1994 Editions

Pasewark
UNDERSTANDING CORPORATE ANNUAL REPORTS – 3.5" VERSION

WE'VE MOVED!
IRWIN
1333 Burr Ridge Pkwy.
Burr Ridge, IL 60521

IRWIN
THE ACCOUNTING PUBLISHER

ACCOUNTING TITLES

AUDITING

NEW	**Auditing** **Fifth Edition**	Thomas D. Hubbard, *University of Nebraska—Lincoln* Johnny R. Johnson, *The American University*

FINANCIAL, MANAGERIAL, AND COST ACCOUNTING

NEW	**Accounting Flows and Disclosures (Financial)**	James J. Benjamin, *Texas A&M University* Arthur J. Francia, *University of Houston* Robert H. Strawser, *Texas A&M University*
NEW	**Financial Accounting** **Third Edition**	Vincent C. Brenner, *Louisiana State University* Ronald M. Copeland, *Northeastern University* Paul E. Dascher, *Drexel University*
NEW	**Financial Accounting Text and Cases**	Paul R. Berney, *Yale School of Organization and Management* Jimie Kusel, *University of Arkansas—Little Rock*
NEW	**Managerial Accounting** **Ninth Edition**	Arthur J. Francia, *University of Houston* Mattie C. Porter, *University of Houston—Clear Lake* Robert H. Strawser, *Texas A&M University* Jerry J. Strawser, *University of Houston*
	Managerial Accounting **Third Edition**	Ronald M. Copeland, *Northeastern University* Paul E. Dascher, *Drexel University* Jerry R. Strawser, *University of Houston*
NEW	**Managerial Accounting** **Third Edition**	Pierre L. Titard, *The University of Alabama—Huntsville*
	Managerial Cost Accounting: Planning and Control **Revised Edition**	William L. Ferrara, *Stetson University* Frank P. Dougherty, *Saint Mary's University* Germain B. Boer, *Vanderbilt University*
	Cost Accounting for Managerial Planning, Decision Making and Control **Fourth Edition**	Woody M. Liao, *University of California—Riverside* James L. Boockholdt, *Samford University*
NEW	**The Executive's Guide to Management Accounting and Control Systems** **Fifth Edition**	**For MBAs and Professional Development** William Rotch, *University of Virginia* Brandt R. Allen, *University of Virginia* C. Ray Smith, *University of Virginia*
NEW	**Income Tax for Accounting Principles** **1993 Edition**	Benny R. Copeland, *University of North Texas* John E. Price, *University of North Texas*
	Practice Case in Financial Statement Preparation and Analysis **The Keller Corporation**	Francis E. Kearns, *Rochester Institute of Technology* Robert P. Derstine, *Villanova University* Ronald J. Heufner, *State University of New York—Buffalo*
	Practice Problem in Financial Accounting **Music Machines, Incorporated** **Revised Edition**	Thomas P. Edmonds, *University of Alabama—Birmingham* Margaret L. Shelton, *University of Houston—Downtown*
	Practice Problem for Intermediate Accounting I **Monroe & Sons, Incorporated** **Third Edition**	Thomas P. Edmonds, *University of Alabama—Birmingham* Margaret L. Shelton, *University of Houston—Downtown*

DAME
PUBLICATIONS, INC.
7800 Bissonnet—Suite 415
Houston, TX 77074
(713) 995-1000 FAX (713) 995-9637

When requesting examination copies, please indicate course title and text(s) currently in use.